



2613

Docket No. 212636US2SRP/5tm

#3

THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Noboru YAMAGUCHI, et al.

GAU: 2613

SERIAL NO: 09/925,567

EXAMINER:

FILED: August 10, 2001

FOR: VIDEO ENCODING APPARATUS AND METHOD AND RECORDING MEDIUM STORING PROGRAMS  
FOR EXECUTING THE METHOD

REQUEST FOR PRIORITY

RECEIVED

ASSISTANT COMMISSIONER FOR PATENTS  
WASHINGTON, D.C. 20231

OCT 31 2001

Technology Center 2600

SIR:

- ☐ Full benefit of the filing date of U.S. Application Serial Number [US App No], filed [US App Dt], is claimed pursuant to the provisions of 35 U.S.C. §120.
- ☐ Full benefit of the filing date of U.S. Provisional Application Serial Number, filed, is claimed pursuant to the provisions of 35 U.S.C. §119(e).
- ☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

COUNTRY

APPLICATION NUMBER

MONTH/DAY/YEAR

JAPAN

2000-245026

August 11, 2000

Certified copies of the corresponding Convention Application(s)

- ☒ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. filed
- ☐ were submitted to the International Bureau in PCT Application Number .  
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and  
(B) Application Serial No.(s)
  - ☐ are submitted herewith
  - ☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.

*Marvin J. Spivak*

Marvin J. Spivak  
Registration No. 24,913

Surinder Sachar  
Registration No. 34,423



22850

Tel. (703) 413-3000  
Fax. (703) 413-2220  
(OSMMN 10/98)



日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office

出願年月日

Date of Application:

2000年 8月11日

出願番号

Application Number:

特願2000-245026

出願人

Applicant(s):

株式会社東芝

RECEIVED

OCT 31 2001

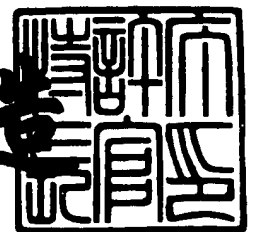
Technology Center 2600

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 7月27日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造



【書類名】 特許願

【整理番号】 A000002646

【提出日】 平成12年 8月11日

【あて先】 特許庁長官 殿

【国際特許分類】 H04N 7/30

【発明の名称】 動画像符号化装置および動画像符号化方法およびプログラムを記録した媒体

【請求項の数】 13

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 山口 昇

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 古川 理恵子

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 菊池 義浩

【特許出願人】

【識別番号】 000003078

【氏名又は名称】 株式会社 東芝

【代理人】

【識別番号】 100058479

【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】	図面	1
【物件名】	要約書	1
【プルーフの要否】	要	

【書類名】 明細書

【発明の名称】 動画像符号化装置および動画像符号化方法およびプログラムを記録した媒体

【特許請求の範囲】

【請求項 1】

入力された動画像信号を解析して、フレーム毎に統計的特徴量を算出する第 1 の特徴量算出手段と、

前記第 1 の特徴量算出手段により算出された特徴量に応じて、動画像を時間的に連続した少なくとも一つのフレームからなるシーンに分割するシーン分割手段と、

第 1 の特徴量算出手段より得られたフレーム毎の特徴量から前記シーン分割手段により分割された各シーン毎の平均的な特徴量を求める第 2 の特徴量算出手段と、

前記シーン分割手段により分割された複数シーンの一部あるいは全てを選択するシーン選択手段と、

前記シーン選択手段により選択されたシーン毎の特徴量をもとにシーン毎の最適なフレームレートと量子化幅の情報を少なくとも含むシーン毎の符号化パラメータを生成する符号化パラメータ生成手段と、

前記符号化パラメータ生成手段により生成された各シーン毎の符号化パラメータに従って前記入力動画像信号を符号化する手段と、

を備えることを特徴とする動画像符号化装置。

【請求項 2】

請求項 1 に記載の動画像符号化装置において、

前記シーン選択手段はユーザによる編集操作により得られる操作情報に従ってシーン選択を実行される構成であることを特徴とする動画像符号化装置。

【請求項 3】

請求項 2 に記載の動画像符号化装置において、

各シーンの内容をユーザに提示するシーン内容提示手段を有することを特徴とする動画像符号化装置。

【請求項 4】

請求項 3 に記載の動画像符号化装置において、

前記シーン内容提示手段は、各シーンの代表フレーム画像あるいは代表フレームの縮小画像（サムネイル画像）をユーザに提示する手段であることを特徴とする動画像符号化装置。

【請求項 5】

請求項 3 に記載の動画像符号化装置において、

前記シーン内容提示手段は、前記第 2 の特徴量算出手段により得られる各シーンの特徴量あるいは特徴を示す記号をユーザに提示する手段であることを特徴とする動画像符号化装置。

【請求項 6】

請求項 3 に記載の動画像符号化装置において、

前記シーン内容提示手段は、各シーンの代表フレーム画像あるいは代表フレームの縮小画像（サムネイル画像）および前記第 2 の特徴量算出手段により得られる各シーンの特徴量あるいは特徴を示す記号をユーザに提示する手段であることを特徴とする動画像符号化装置。

【請求項 7】

符号化された動画像のビットストリームの全てあるいは一部を解析して、フレーム毎に統計的特徴量を算出する第 1 の特徴量算出ステップと、

この第 1 の特徴量算出ステップにより算出された特徴量に応じて、動画像を時間的に連続した少なくとも一つのフレームからなるシーンに分割するシーン分割ステップと、

このシーン分割ステップにより分割された各シーン毎の平均的な特徴量を前記第 1 の特徴量算出ステップより得られたフレーム毎の特徴量から求める第 2 の特徴量算出ステップと、

前記シーン分割ステップにより分割された複数シーンの一部あるいは全てを選択するシーン選択ステップと、

シーン選択ステップにより選択されたシーン毎の特徴量から各シーン毎の最適なフレームレートと量子化幅の情報を少なくとも含むシーン毎の符号化パラメー

タを生成する符号化パラメータ生成ステップと、

この符号化パラメータ生成ステップにより生成された符号化パラメータを用いて前記入力ビットストリームから再生された動画像信号を再度符号化する符号化ステップと、

を備えることを特徴とする動画像符号化方法。

【請求項 8】

請求項 7 に記載の動画像符号化方法において、

前記シーン選択ステップはユーザによる編集操作により実行されることを特徴とする動画像符号化方法。

【請求項 9】

請求項 8 に記載の動画像符号化方法において、各シーンの内容をユーザに提示するシーン内容提示ステップを有することを特徴とする動画像符号化方法。

【請求項 10】

請求項 9 に記載の動画像符号化方法において、

シーン内容提示ステップは、各シーンの代表フレーム画像あるいは代表フレームの縮小画像（サムネール画像）をユーザに提示するものであることを特徴とする動画像符号化方法。

【請求項 11】

請求項 9 に記載の動画像符号化方法において、

前記シーン内容提示ステップは、前記第 2 の特徴量算出ステップにより得られる各シーンの特徴量あるいは特徴を示す記号をユーザに提示するものであることを特徴とする動画像符号化方法。

【請求項 12】

請求項 9 に記載の動画像符号化方法において、

前記シーン内容提示ステップは、各シーンの代表フレーム画像あるいは代表フレームの縮小画像（サムネール画像）および前記第 2 の特徴量算出ステップにより得られる各シーンの特徴量あるいは特徴を示す記号をユーザに提示するものであることを特徴とする動画像符号化方法。

【請求項 13】

符号化された動画像のビットストリームの全てあるいは一部を解析して、フレーム毎に統計的特徴量を算出する第1の特徴量算出ステップと、

この第1の特徴量算出ステップにより算出された特徴量に応じて、動画像を時間的に連続した少なくとも一つのフレームからなるシーンに分割するシーン分割ステップと、

このシーン分割ステップにより分割された各シーン毎の平均的な特徴量を前記第1の特徴量算出ステップより得られたフレーム毎の特徴量から求める第2の特徴量算出ステップと、

前記シーン分割ステップにより分割された複数シーンの一部あるいは全てを選択するシーン選択ステップと、

シーン選択ステップにより選択されたシーン毎の特徴量から各シーン毎の最適なフレームレートと量子化幅の情報を少なくとも含むシーン毎の符号化パラメータを生成する符号化パラメータ生成ステップと、

この符号化パラメータ生成ステップにより生成された符号化パラメータを用いて前記入力ビットストリームから再生された動画像信号を再度符号化する符号化ステップと、

からなるコンピュータ読み取りおよび実行可能なプログラムを記録した媒体。

#### 【発明の詳細な説明】

##### 【0001】

#### 【発明の属する技術分野】

本発明は、インターネットなどによる動画像伝送システムや画像データベースシステムに使用されるMPEG方式などの動画像圧縮符号化装置技術に係り、特に2パス符号化と呼ばれる手法によりシーンの内容に応じた符号化パラメータに従った符号化を行う動画像符号化装置および動画像符号化方法に関する。

##### 【0002】

#### 【従来の技術】

動画像符号化の国際標準方式としてMPEG1 (Motion Picture Experts Group-1), MPEG2 (Motion Picture Experts Group-2)およびMPEG4 (Motion Picture Experts Group-4)があり、実用化されている。そして、これらにおいて

は、符号化の基本方式としてMC+DCT方式が採用されている。以下、参考文献（文献1：三木 編著、“MPEG-4のすべて”第3章、工業調査会、1998）を参照して、MPEG4検証(Verification)モデルに従って説明する。

### 【0003】

#### <MC+DCT方式の概要>

図13を用いてMC+DCT方式の符号化法を説明する。MC（動き補償予測）+DCT（Discrete Cosine Transform；離散コサイン変換（直交変換））方式は、図13に示されるように、差分値算出部101、動き補償予測部（MC）102、フレーム内／フレーム間（Intra / inter）切り替え部103、フレームメモリ（FM）104、動きベクトル検出部（ME）105、離散コサイン変換部（DCT）106、量子化部（Q）107、可変長符号化部（VLC）108、逆量子化部（IQ）109、逆離散コサイン変換部（IDCT）110、加算部111、出力バッファ（Buffer）112、レート制御部（Rate Control）113とから構成される。

### 【0004】

差分値算出部101において、画像信号入力線11を介して供給される画像信号と、信号線12を介して供給される動き補償予測信号との差分が計算され、この差分信号は信号線13を介して離散コサイン変換部106に供給される。

### 【0005】

信号線12には、フレーム内符号化（Inter）モードの場合には、動き補償予測部102で生成される動き補償予測信号がフレーム内／フレーム間切り替え部103を介して供給され、一方、フレーム間符号化（Intra）モードの場合には、信号が供給されない。つまり、フレーム間符号化（Intra）モードの場合には、信号線13には差分信号ではなく画像信号入力線11の信号がそのまま供給される。

### 【0006】

Intra / Inter モードの切り替えは、後述するように動きベクトル検出部105で判定され、信号線14を介してフレーム内／フレーム間切り替え部103に供給される。動き補償予測部102で生成される動き補償予測信号は、フレーム

メモリ 1 0 4 に蓄積されている既に符号化済みのフレームの信号から、動きベクトル検出部 1 0 5 で検出された動きベクトル情報にしたがって生成される。

【 0 0 0 7 】

離散コサイン変換部 1 0 6 では、信号線 1 3 を介して供給される信号を離散コサイン変換し、その変換結果としての D C T 変換係数（離散コサイン変換係数）出力を、量子化部 1 0 7 に供給する。量子化部 1 0 7 ではこの D C T 変換係数を所定の量子化幅で量子化する。

【 0 0 0 8 】

後述するように、量子化部 1 0 7 と逆量子化部 1 0 9 は、与えられる量子化パラメータに従って量子化幅、逆量子化幅が定まる構成であり、量子化部 1 0 7 で量子化された D C T 変換係数は、可変長符号化処理するために可変長符号化部 1 0 8 に供給されると共に、また、逆量子化部 1 0 9 にも供給されて逆量子化される。そして、逆量子化部 1 0 9 により逆量子化された変換係数は、逆離散コサイン変換部 1 1 0 に供給されて信号線 1 3 に対する再生信号が生成され、この生成された再生信号は加算部 1 1 1 に供給される。

【 0 0 0 9 】

加算部 1 1 1 では、逆離散コサイン変換部 1 1 0 から供給される信号と、信号線 1 2 を介して供給される信号とを加算して画像信号を再生した後、フレームメモリ 1 0 4 に蓄積する。

【 0 0 1 0 】

一方、可変長符号化部 1 0 8 では、量子化部 1 0 7 で量子化されて与えられた D C T 変換係数や、動きベクトル情報（図示せず）などを受けて、これらを可変長符号化処理した後、多重化してビットストリームを生成し、出力バッファ 1 1 2 に供給する。

【 0 0 1 1 】

出力バッファ 1 1 2 は保持しているビットストリームの蓄積量情報を出力できる構成であり、ビットストリームを供給された当該出力バッファ 1 1 2 ではこれを一時保持し、ネットワークや蓄積媒体へ、その特性に応じてレートを制御しながら当該ビットストリームを信号線 1 5 を介して出力する。

## 【 0 0 1 2 】

レート制御部 1 1 3 では、信号線 1 6 を介して供給される出力バッファ 1 1 2 のビットストリーム蓄積量に応じて、量子化パラメータを決定し、信号線 1 7 を介して量子化部 1 0 7 と逆量子化部 1 0 9 に供給する。

## 【 0 0 1 3 】

ここで、出力バッファ 1 1 2 内のビットストリーム蓄積量が大きくなってきた場合は、量子化パラメータを大きくして量子化部 1 0 7 での発生符号量を少なくし、出力バッファ 1 1 2 内の蓄積量が小さくなってきた場合は、量子化パラメータを小さくすることで、発生符号量が一定になるように制御される。

## 【 0 0 1 4 】

このように、レート制御は量子化幅を変更することにより行うが、画像の変化の激しいシーンでは、符号量が増大することから、それにあわせて量子化幅を大きくすることとなるので、その結果、量子化幅が適切でなくなることもあり、このような場合に、画像が歪むなど、画質劣化の問題を引き起こす。このことを少し詳しく説明する。

## 【 0 0 1 5 】

M P E G 方式に基づく従来の動画像符号化方式では、圧縮動画像データを伝送レートが規定されている伝送路によって伝送したり、記録容量に限りがある蓄積媒体に記録するために、出力される符号化ビットストリームのビットレートが指定された値となるようにフレームレートや量子化幅などの符号化パラメータを設定して符号化を行うというレート制御という処理が行われている。

## 【 0 0 1 6 】

多くのレート制御では、前フレームの発生符号量に応じて次のフレームまでの間隔と次フレームの量子化幅を決める方法がとられている。

## 【 0 0 1 7 】

そのため、発生符号量が増大することとなる画面の動きが大きいシーンになると、符号量増大に対応するめに量子化幅を大きくする方向に制御されることから、画質が急に劣化する。図 1 4、図 1 5 は従来のレート制御について示す図である。図 1 4 の ( a ) に示す [ I ] のグラフは、ビットレートの推移の例を示して

おり、401が目標のビットレート（ビットレート目標値）、402が実際のビットレートを示している。時刻 $t_{11}$ から $t_{12}$ の期間に動きの激しいシーンが発生してこれに伴って発生ビットレートが大きくなると、バッファが溢れるので、フレームレートを低くしようとする制御がかかる。これが図14（b）に示す[II]のグラフであり、ビットレートが急増した時刻 $t_{11}$ において、フレームレートを抑えるべくレート制御が働く結果、実際のフレームレート404が急激に低下している様子が示されている。

#### 【0018】

そして、一方、レート制御では、予め設定されたフレームスキップ閾値のバッファサイズと現時点でのバッファレベルとの差（余裕度）を基に、フレームレートを決め（図15（a）参照）、現時点でのバッファが閾値よりも小さい時には一定のフレームレートで符号化を行い、現時点でのバッファが閾値を上回るとフレームレートを下げるように制御する。

#### 【0019】

このような制御を行う結果、発生符号量が大きいフレーム $f_m$ ではフレームレートが低くなって、例えば図15（b）に示すように、 $f_{21}$ 、 $f_{22}$ と今まで等間隔で来た発生フレームが、 $f_{22}$ と $f_{23}$ の間のように、フレームの間隔が広がるという現象が生じることとなる。つまり、フレームのスキップが生じるわけである。

#### 【0020】

これは従来のレート制御が、画像内容に無関係に次フレームの符号量を定めていることに起因するものであって、そのため、画面の動きが大きくなるシーンでは、このように、フレーム間隔が広くなり過ぎて画像の動きが不自然になったり、量子化幅が適切でないために画像が歪んで見づらさを感じさせることがあった。

#### 【0021】

そこで、このような問題点を解決する必要があるが、そのための手法として、既にいくつかのものが知られている。その中の2パス符号化と呼ばれる方法によりレート制御を行う方式（文献2：特開平08-186821号公報参照）を

除くと、他の多くは符号量の変化のみに着目する方法が主体であり、画像内容と符号量の関係を考慮したものは、例えばフェードイン・フェードアウトなど（文献3：特開平10-336641号公報参照）の特別な場合に限られていた。

【0022】

そこで、本件発明者らは、2パス符号化の一つとして、解析されたシーン内容に応じてビットレートを配分し、全体的なビットレートが予め指定されたビットレートを満たすように効率良く符号化パラメータを配分する動画像符号化方法および装置を提案した（文献4：特願2000-053823号）。

【0023】

また、シーンの内容を解析して、シーン毎に映像に対して撮影者の意図を代表する見出しを自動的に作成して提示することで、一般の人でも容易に編集が可能となる動画像編集システムの提案もある（文献5：堀他，“映像解析技術年利用した映像メディアのためのGUI”，ヒューマンインタフェース72-7pp. 37-42，1997）。但し、この編集システムではシーン内容を符号化に反映することは無かった。

【0024】

一方、蓄積メディア向けに符号化データを作成する場合に、編集装置により予め動画像を編集して符号化している。従来、編集操作の結果を符号化に利用したとしても、編集時のカット点を考慮する程度であった。

【0025】

【発明が解決しようとする課題】

上述のように、従来の動画像符号化装置では、画像の内容に無関係にフレームレートや量子化幅が決められていたために、オブジェクトの動きが激しいシーンにおいて急激にフレームレートが低下したり、量子化幅が適切でないために画像が歪むなど、画質劣化が目立ち易いという問題点があった。

【0026】

また、動画像信号をパソコンなどを使用してカット&ペースト等を行い、所望の映像展開となるように編集して動画を完成させるようにするが、この編集操作で、シーンの内容を把握していても、その情報を動画像信号の符号化時に活かす

仕組みを持っていなかった。そのため、ビットレートの配分なども無駄の多いものとなっていた。

【0027】

そこで、本発明の目的とするところは、動画像を解析し、シーンの内容を編集操作に活用すると共に、シーンの内容に応じてビットレートを適正に配分し、全体的なビットレートが予め指定されたビットレートを満たすように効率良く符号化パラメータを配分することができるようにした動画像符号化方法および動画像編集方法を提供することにある。

【0028】

【課題を解決するための手段】

上記の目的を達成するため、本発明は、画像信号を符号化するにあたり、第1パス（最適化準備モード）でパラメータの最適化を実施し、第2パス（実行モード）ではこの最適化したパラメータを用いて符号化処理をする。

【0029】

すなわち、本発明においては、まず入力動画像信号を時間的に連続した少なくとも一つのフレームからなるシーンに分割して、各シーン毎に統計的特徴量（シーン毎の全フレームについての、フレーム内のマクロブロックの動きベクトルと動き補償残差、輝度値の平均・分散）を算出し、この統計的特徴量に基づいてシーンの内容を推定する。シーンの内容は編集操作にも活用され、編集によりシーンのカットアンドペーストが発生してもシーン毎の統計的特徴量の相対的な関係を利用して目標ビットレートに対し、最適な符号化パラメータを決定する。そして、この符号化パラメータを用いて入力画像信号を符号化するようにし、これによって、同じデータサイズでも視覚的に見易い復号画像を得ることを基本的な特徴とする。

【0030】

ここで、統計的特徴量は、例えば、入力動画像信号の各フレーム内に存在する動きベクトルや輝度値をシーン毎に集計することにより算出される。加えて、特徴量から入力動画像信号を得るときに使用したカメラの動きおよび画像内のオブジェクトの動きを推定した結果を用いて、それらを符号化パラメータに反映させ

る。また、マクロブロック毎に輝度値の分布を調べることで、モスキートノイズが発生しやすいマクロブロックやオブジェクトのエッジが存在するマクロブロックの量子化幅を他のマクロブロックに比して相対的に小さくし、画質の向上を図る。

#### 【0031】

そして、第2パス目における符号化では、前記算出されたシーン毎に適切なビットレートとフレームレートを与えることで、従来のレート制御機構を大きく変更することなくシーン内容に応じた符号化が行える。

#### 【0032】

上記のような2パスの手法を用いて、目標符号量と同じデータサイズで良好な復号画像を得る符号化を実現することができる。

#### 【0033】

##### 【発明の実施の形態】

本発明は、画像信号を符号化するにあたり、第1パス（最適化準備モード）でパラメータの最適化を実施し、第2パス（実行モード）ではこの最適化したパラメータを用いて符号化処理をする。具体的には、まず入力動画像信号を時間的に連続した少なくとも一つのフレームからなるシーンに分割して、各シーン毎に統計的特徴量を算出し、この統計的特徴量に基づいてシーンの内容を推定する。シーンの内容は編集操作にも活用され、編集によりシーンのカットアンドペーストが発生してもシーン毎の統計的特徴量の相対的な関係を利用して目標ビットレートに対し、最適な符号化パラメータを決定する。これが第1パスの処理である。そして、第2パスでは、この符号化パラメータを用いて入力画像信号を符号化するようにする。これにより、同じデータサイズでも視覚的に見易い復号画像を得ることができるようにする。

#### 【0034】

以下、図面を参照して本発明の実施の形態を説明する。

#### 【0035】

##### （第1の実施形態）

図1は本発明の一実施形態に係る動画像編集・符号化装置の構成を示すブロッ

ク図である。図において、100はエンコーダ、120はサイズ変換部、200はソースデータ、210はデコーダ、220は特徴量計算部、230は構造化情報蓄積部、240は構造化情報提示装置、250は最適パラメータ計算部、260は最適パラメータ蓄積部である。

## 【0036】

これらのうち、エンコーダ100は、サイズ変換部120を介して与えられる動画像信号を符号化して出力するためのものであって、最適パラメータ蓄積部260の蓄積しているパラメータ（各シーン別の最適なフレームレートと量子化幅の情報）を用いて動画像信号の符号化をするものである。

## 【0037】

デコーダ210は、入力されるソースデータ200のフォーマットに対応しており、信号線20を介して入力されるソースデータ200を復号処理して元の動画像信号を再生するものである。このデコーダ210で再生された動画像信号は信号線21を介して特徴量計算部220とサイズ変換部120とに供給されるようになっている。

## 【0038】

なお、ソースデータ200は、同じ信号を繰り返し複数回再生可能なデジタルVTRやDVDシステムのようなビデオ記録再生装置に記録されている動画像のデータである。

## 【0039】

特徴量計算部220は、デコーダ210から与えられた動画像信号についてシーン分割を行うと同時に動画像信号の各フレームに対し画像特徴量を計算する機能を有している。ここで云う画像特徴量とは、例えば、“動きベクトルの数”、“分布”、“ノルムの大きさ”、“動き補償後の残差”、“輝度・色差の分散”等である。特徴量計算部220はこの算出した特徴量およびシーンの代表フレーム画像を前記分割されたシーン毎に集計し、信号線22を介して構造化情報蓄積部230に供給するように構成されている。

## 【0040】

構造化情報蓄積部230は、各シーン毎の代表フレーム画像や特徴量の情報を

シーン毎に構造化した情報として蓄積するものである。なお、代表フレーム画像のサイズが大きい場合には、その縮小画像（サムネール画像）を代わりに蓄積しても良い。

#### 【0041】

構造化情報提示装置240は、キーボード等の入力装置およびマウス等のポインティングデバイスを少なくとも有し、また、ディスプレイを有したマンマシンインターフェースであって、入力装置を用いての編集操作をはじめとする各種操作入力や指示入力をしたり、構造化情報蓄積部230に蓄積されている各シーン毎の代表フレーム画像や特徴量を信号線23を介して受けることにより、これらを図2のような提示形態でディスプレイに表示して、動画像信号の内容をユーザに提示するものである。

#### 【0042】

なお、本発明システムでは、第2パス目の処理においては、信号線21を介して供給される動画像信号は、信号線24を介して構造化情報呈示装置240から供給される編集情報対応に編集されたソースデータを、デコーダ210で再生して得た動画像信号である。

#### 【0043】

サイズ変換部120は、信号線21を介して供給される動画像信号の画面サイズと、エンコーダ100で符号化して出力する動画像信号の画面サイズが異なる場合に、画面サイズを変換する処理をするものであって、エンコーダ100はこのサイズ変換部120の出力を信号線11を介して受けて符号化処理をするものである。

#### 【0044】

また、最適パラメータ計算部250は、構造化情報蓄積部230から与えられる特徴量の情報の供給を信号線25を介して受け、各シーンに対して最適なフレームレートと量子化幅を算出するものであり、構造化情報蓄積部230から読み出す特徴量の情報は、信号線24を介して供給される構造化情報呈示装置240からの編集情報に従って、構造化情報蓄積部230が該当するシーンの特徴量の情報を読み出し、供給する構成である。

## 【 0 0 4 5 】

また、最適パラメータ蓄積部 2 6 0 はこの最適パラメータ計算部 2 5 0 の計算した各シーン別の最適なフレームレートと量子化幅の情報を蓄積するためのものである。

## 【 0 0 4 6 】

次に、このような構成の本システムの作用を説明する。本発明システムでは、最初に第 1 パスの処理（最適化準備モード）を実施し、次に第 2 パスの処理（実行モード）を実施する 2 段階処理で実施する方式である。そのため、本システムでは、何回でも繰り返し同一の動画像信号を再生して供給できるデジタル V T R や D V D システムのようなビデオ記録再生装置を用い、このビデオ記録再生装置に記録されているデータを再生してこれをソースデータ 2 0 0 として、信号線 2 0 よりデコーダ 2 1 0 に供給する。

## 【 0 0 4 7 】

このビデオ記録再生装置からソースデータ 2 0 0 を受けたデコーダ 2 1 0 は、当該ソースデータを復号処理し、動画像信号として出力する。そして、このデコーダ 2 1 0 で再生された動画像信号は、第 1 パスにおいては信号線 2 1 を介して特徴量計算部 2 2 0 に供給される。

## 【 0 0 4 8 】

特徴量計算部 2 2 0 では、この動画像信号を用いてまず動画像信号のシーン分割を行う。同時に動画像信号の各フレームに対し、画像特徴量を計算する。ここで云う画像特徴量とは、例えば動きベクトルの数、分布、ノルムの大きさ、動き補償後の残差、輝度・色差の分散等である。

## 【 0 0 4 9 】

そして、この特徴量算出部 2 2 0 では、シーンの代表フレーム画像および算出したこのような特徴量を、前記分割されたシーン毎に集計し、信号線 2 2 を介して構造化情報蓄積部 2 3 0 に供給する。

## 【 0 0 5 0 】

そして、構造化情報蓄積部 2 3 0 は、これらの情報を蓄積する。その結果、第 1 パスにおいては、構造化情報蓄積部 2 3 0 には、供給動画像信号を解析して得

たシーン毎に構造化した情報が蓄積されることになる。なお、分割したシーン毎の代表フレーム画像を蓄積するに当たり、当該代表フレーム画像のサイズが大きい場合には、その縮小画像（サムネール画像）を代わりに蓄積しても良い。

#### 【0051】

このようにして構造化情報蓄積部230に動画像信号の各シーン毎の特徴量、代表フレーム画像が蓄積されると、次に、構造化情報蓄積部230は蓄積してある各シーン毎の代表フレーム画像や特徴量を読み出し、信号線23を介して構造化情報提示装置240に供給する。これを受けた構造化情報提示装置240では、図2のような提示形態により、動画像信号の内容をユーザに提示する。

#### 【0052】

図2の例は、前述の文献5で開示されている例であり、各シーン毎の代表フレーム画像f a, f b, f c, f dと、これらそれぞれの画像f a, f b, f c, f dの動きの内容情報m a, m b, m c, m dを、画面に表示することでユーザに提示し、以て各シーンの内容をユーザに容易に想起させることができるようにしている。

#### 【0053】

構造化情報提示装置240には、代表フレーム画像を画面上でカット&ペーストしたり、ドラッグ&ドロップ操作することなどで、位置の移動やシーンの削除、コピーなどといった編集操作を自由に行うことができる動画像の編集機能を備えており、従って、上述したように、動画像信号の代表フレーム画像と構造化情報をユーザに提示することで、ユーザは動画像信号の内容を容易に把握できるようになり、図3に示される如く、シーンのカット&ペーストなどの編集操作が容易に行えるようになる。もちろん、複数の動画像信号の構造化情報をユーザに提示し、編集することも可能である。

#### 【0054】

図3の例は、オリジナルでは（a）の如く配置されていた図2の表示状態のものについて、代表フレームf cをカットし、代表フレームf cとf dを入れ替えて代表フレームf aで代表されるシーンの次に代表フレームf dで代表されるシーンが入り、その次に代表フレームf bで代表されるシーンに移っていくといっ

た内容に編集されたことを示している（図 3（b））。

【 0 0 5 5 】

例えば、このようにしてユーザの編集操作により編集された編集情報は、信号線 2 4 を介して構造化情報蓄積部 2 3 0 とソースデータ 2 0 0 に供給される。ここで編集情報とは、どのシーンが選択されたかの情報と、選択されたシーンのソースデータ 2 0 0 におけるタイムスタンプや、編集後のシーン配置などの情報である。

【 0 0 5 6 】

ユーザが構造化情報提示装置 2 4 0 を使用して上述の如き編集を行うと、その情報は編集情報として信号線 2 4 を介して構造化情報蓄積部 2 3 0 に供給され、当該構造化情報蓄積部 2 3 0 はこの編集情報を蓄積すると同時に、最適パラメータ計算部 2 5 0 にも与える。

【 0 0 5 7 】

最適パラメータ計算部 2 5 0 では、供給された編集情報に従って、構造化情報蓄積部 2 3 0 に蓄積されている該当するシーンの特徴量の情報の供給を受け、各シーンに対して最適なフレームレートと量子化幅を算出して、最適パラメータ蓄積部 2 6 0 に与える。これにより最適パラメータ蓄積部 2 6 0 には各シーン毎に、最適なフレームレートと量子化幅の情報が蓄積されることになる。

【 0 0 5 8 】

最適パラメータ計算部 2 5 0 の具体的例を図 4 を用いて説明する。

【 0 0 5 9 】

＜最適パラメータ計算部 2 5 0 の構成＞

最適パラメータ計算部 2 5 0 は、ユーザが構造化情報提示装置 2 4 0 を編集操作することにより、当該構造化情報提示装置 2 4 0 から与えられる編集情報に従って、構造化情報蓄積部 2 3 0 から該当するシーンの特徴量を受け、各シーンに対して最適なフレームレートと量子化幅を算出するものであるが、当該最適パラメータ計算部 2 5 0 は、図 4 に示す如く、符号化パラメータ生成部 2 5 1、発生符号量子予測部 2 5 2、符号化パラメータ修正部 2 5 3 より構成される。

## 【0 0 6 0】

これらのうち、符号化パラメータ生成部 2 5 1 は、構造化情報蓄積部 2 3 0 から受けた特徴量をもとに、各シーンの特徴量の相対的な関係から、各シーンに適切なフレームレートと量子化幅を算出するものであり、発生符号量予測部 2 5 2 は、この符号化パラメータ生成部 2 5 1 により算出されたフレームレートと量子化幅で動画像信号を符号化した際の発生符号量を予測するものである。

## 【0 0 6 1】

また、符号化パラメータ修正部 2 5 3 は、パラメータの修正を行うためのものであって、予測した符号量がユーザ設定の符号量を満たすようにパラメータの修正を行うことで最適なパラメータを求めるものである。

## 【0 0 6 2】

このような構成の 最適パラメータ計算部 2 5 0 においては、信号線 2 5 を介して構造化情報蓄積部 2 3 0 から供給されたシーン毎の特徴量は、符号化パラメータ生成部 2 5 1 で各シーンの特徴量の相対的な関係から、各シーンに適切なフレームレートと量子化幅を算出する。そして、発生符号量予測部 2 5 2 では、これらを入力とし、この算出されたフレームレートと量子化幅で動画像信号の符号化を行った際の発生符号量を予測する。

## 【0 0 6 3】

その際に、予測した発生符号量がユーザが設定した目標符号量 2 5 4 と大きく異なる場合には、予測した符号量がユーザが設定した符号量を満たすように符号化パラメータ修正部 2 5 3 においてパラメータの修正を行うことで最適なパラメータを求めている。

## 【0 0 6 4】

以上、動画像信号を再生し、シーン毎に特徴量の情報と代表フレーム画像を求め、保存し、これを用いて動画像信号の編集操作をすると、その編集情報に従って対応するシーンの特徴量を読み出し、これを用いて各シーン毎に最適なフレームレートと量子化幅を算出してその情報をパラメータとして蓄積するというのが 1 パス目の処理である。

## 【0 0 6 5】

このような第1パスの処理が終わると、ユーザは構造化情報提示装置240を操作して、モードを実行モード、すなわち、第2パスでの処理モードに切り替える。すると、構造化情報提示装置240は動画像信号を最適パラメータ蓄積部260に蓄積された各シーン毎の最適なフレームレートと量子化幅の情報を用いてエンコーダ100により符号化処理させるべくシステムを駆動させるコマンドを発生する。

## 【0 0 6 6】

これにより、システムは第2パスの処理（実行モード）を開始する。

## 【0 0 6 7】

第2パス目の処理においては、信号線21を介して供給される動画像信号は、信号線24を介して供給される編集情報により、ソースデータ200を編集した編集処理済みのソースデータを、デコーダ210で再生した動画像信号である。

## 【0 0 6 8】

この動画像信号は、エンコーダ100に送られ、その各シーン毎に、最適パラメータ蓄積部260に蓄積されているそのシーン対応の最適パラメータを用いて符号化処理される。その結果、エンコーダ100からは、シーンの内容に応じて適切に符号量が配分されたビットストリーム15が出力される。

## 【0 0 6 9】

このように、第2パス目の処理は、信号線21を介して供給される動画像信号をエンコーダ100で符号化するが、その符号化には前記最適パラメータ蓄積部260に蓄積されている最適パラメータを用いて符号化することで、シーンの内容に応じて適切に符号量が配分されたビットストリームにするとする処理である。この結果、動画像を解析し、シーンの内容を編集操作に活用すると共に、シーンの内容に応じてビットレートを配分し、全体的なビットレートが予め指定されたビットレートを満たすように効率良く符号化パラメータを配分する動画像符号化が行えることになり、スキップが生じることがなく、また、同じデータサイズでも視覚的に見易い復号画像を得ることができる符号化方法を提供できる。

## 【0 0 7 0】

なお、第2パスにおいて、信号線21を介して供給される動画像信号の画面サイズと、エンコーダ100で符号化する画面サイズが異なる場合には、サイズ変換部120において画面サイズを変換した後、その動画像信号を信号線11を介してエンコーダ100に供給する。これにより、画面サイズの不整合による問題も生じなくなる。

#### 【0071】

次に、本実施形態のシステムにおける特徴量計算部220での個々の処理についてさらに詳しく説明する。画像特徴量を算出処理する特徴量計算部220での画像特徴量算出処理の対象は、入力される動画像信号に対してのシーン分割の処理、入力動画像信号の全フレームについて、フレーム内のマクロブロックの動きベクトルと動き補償残差、輝度値の平均・分散等を計算する処理である。そして、これによって得られたシーン毎の入力動画像信号の全フレームについての、フレーム内のマクロブロックの動きベクトルと動き補償残差、輝度値の平均・分散等が画像特徴量である。

#### 【0072】

##### <特徴量計算部でのシーン分割処理>

特徴量計算部220において、入力動画像信号21は、隣接するフレーム間の差分によりフラッシュフレームや雑音フレームなどのフレームを除いて複数のシーンに分割される。ここで、フラッシュフレームとは、例えば、ニュース番組でのインタビューシーンで、フラッシュ（ストロボ）が発光した瞬間のように、輝度が急激に高くなるフレームである。また、雑音フレームとはカメラの振れ等により画像が大きく劣化したフレームである。

#### 【0073】

例えば、シーン分割は以下の様に行う。

#### 【0074】

図5のように隣接した*i*番目のフレームと(*i*+1)番目のフレームの間の差分値が、予め定められた閾値を超え、かつ、*i*番目のフレームと(*i*+2)番目のフレームの間の差分値も同様に閾値を超えているならば、(*i*+1)番目のフレームはシーンの区切りと判定する。

## 【0075】

i 番目のフレームと (i + 1) 番目のフレームの間の差分値が予め定められたしきい値を超えていても、i 番目のフレームと (i + 2) 番目のフレームの間の差分値がしきい値を超えていなければ、(i + 1) 番目のフレームはシーンの区切りとしない。

## 【0076】

## ＜特徴量計算部での動きベクトルの計算＞

特徴量計算部 220 においては、上述のようなシーン分割の処理の他に、入力動画信号 21 の全フレームについて、フレーム内のマクロブロックの動きベクトルと動き補償残差、輝度値の平均・分散等を計算する。なお、特徴量の計算は全フレームに対してでも良いし、画像の性質を解析できる範囲で数フレームおきに計算しても良い。

## 【0077】

i 番目のフレームについての動領域のマクロブロックの数を “M v N u m ( i ) ”、動き補償残差を “M e S a d ( i ) ”、輝度値の分散を “Y v a r ( i ) ” とする。ここで、動領域とは 1 フレーム中で前フレームからの動きベクトル ≠ 0 であるマクロブロックの領域を指す。j 番目のシーンに対して、そのシーンに含まれるフレームすべての M v N u m ( i )、M e S a d ( i )、Y v a r ( i ) の平均値をそれぞれ M v n u m \_ j、M e S a d \_ j、Y v a r \_ j とし、それらを j 番目のシーンの特徴量の代表値とする。

## 【0078】

## ＜特徴量計算部でのシーン分類処理＞

さらに本実施形態では、特徴量計算部 220 においては、動きベクトルを用いて次のようなシーンの分類を行い、シーン内容を推定する。

## 【0079】

すなわち、各々のフレームに対する動きベクトルを算出した後、動きベクトルの分布を調べ、シーンを分類する。具体的には、まずフレーム中の動きベクトルの分布を計算して、各フレームが図 6 ( a ) ～ ( e ) に示す 5 つのタイプのいずれに属するかを調べる。

## 【0080】

タイプ〔1〕： 図6（a）に示すタイプであって、フレーム中に動きベクトルがほとんど存在しないタイプ（動領域のマクロブロック数が $M_{min}$ 以下）。

## 【0081】

タイプ〔2〕： 図6（b）のタイプであって、同じ向き・大きさの動きベクトルが画面全体に分布しているタイプ（動領域のマクロブロック数が $M_{max}$ 以上で大きさと方向がある範囲内にある）。

## 【0082】

タイプ〔3〕： 図6（c）に示すタイプであって、フレーム中で特定の部分にだけ動きベクトルが現れているタイプ（動領域のマクロブロックの位置が特定の部分に集中している）

タイプ〔4〕： 図6（d）に示すタイプであって、フレーム中に放射状に動きベクトルが分布しているタイプ。

## 【0083】

タイプ〔5〕： 図6（e）に示すタイプであって、フレーム中の動きベクトルの数が多く、方向も不揃いのタイプ。

## 【0084】

これらタイプ〔1〕～タイプ〔5〕のパターンは、いずれも処理対象となる動画信号を得るときに使用したカメラや、撮影された画像内のオブジェクトの動きと密接に関係している。すなわち、タイプ〔1〕のパターンではカメラもオブジェクトもいずれも静止している状態にある。また、タイプ〔2〕のパターンは、カメラの平行移動時、タイプ〔3〕のパターンは静止している背景の中でオブジェクトが動いている場合に得られる。また、タイプ〔4〕のパターンは、カメラがズームングを行っている場合に得られる。また、タイプ〔5〕のパターンはカメラとオブジェクトが共に動いている場合に得られる。

## 【0085】

以上のように各フレーム毎に分類した結果をシーン毎にまとめ、シーンが図6（a）～（e）のどのタイプに属するかを判定する。判定されたシーンのタイプと前記算出された特徴量を用いて後述の符号化パラメータ生成部で符号化パラメ

ータであるフレームレートとビットレートをシーン毎に決定する。

【0086】

このようにして、特徴量計算部220においては、動きベクトルを用いてシーンの分類を行い、シーン内容を推定する。

【0087】

次に最適パラメータ計算部250での構成要素の一つである符号化パラメータ生成部251における符号化パラメータ生成に当たっての個々の処理について詳しく説明する。

【0088】

符号化パラメータ生成部251では、(i) フレームレート算出処理、(ii) 量子化幅算出処理、(iii) フレームレートと量子化幅の修正処理、(iv) マクロブロック毎の量子化幅の設定処理、の4種の処理を実施する。そして、これにより、フレームレート、量子化幅およびマクロブロック毎の量子化幅、といった符号化パラメータを生成する。

【0089】

＜符号化パラメータ生成部でのフレームレート算出処理＞

符号化パラメータ生成部251では、まずフレームレートを算出する。

【0090】

このとき、前述の特徴量計算部220においては、シーン毎の特徴量の代表値を既に算出しているものとする。これに対して、j番目のシーンのフレームレートFR(j)を

$$FR(j) = a \times MVnum\_j + b + w\_FR \quad \dots \text{式(1)}$$

により算出する。ただし、MVnum\_jはj番目のシーンの代表値、a、bはユーザが指定したビットレートと画像サイズに関係する係数、w\_FRは後述する重みパラメータである。式(1)は、動きベクトルの代表値MVnum\_jが大きくなるほどフレームレートFR(j)が高くなることを意味している。すなわち、動きの大きいシーンほどフレームレートが高くなる。

## 【0091】

また、動きベクトルの代表値 $MVnum\_j$ は、前述のフレーム中の動きベクトルの数の他にフレーム中の動きベクトルの大きさの絶対値和、密度なども用いることもある。

## 【0092】

以上が、符号化パラメータ生成部251におけるフレームレート算出処理である。

## 【0093】

＜符号化パラメータ生成部での量子化幅の算出処理＞

符号化パラメータ生成部251では、量子化幅の算出に当たり、各々のシーンに対するフレームレートを算出した後、次に各々のシーンに対する量子化幅を計算する。 $j$ 番目のシーンに対する量子化幅 $Qp(j)$ はフレームレート $FR(j)$ と同様、シーンの動きベクトルの代表値 $MVnum\_j$ を用いて以下の式で算出する。

## 【0094】

$$Qp(j) = c \times MVnum\_j + d + w\_Qp \quad \dots \text{式(2)}$$

ここで、 $c$ 、 $d$ はユーザが指定したビットレートと画像サイズに対する係数であり、 $w\_Qp$ は後述する重みパラメータである。

## 【0095】

式(2)は動きベクトルの代表値 $MVnum\_j$ が大きくなるほど量子化幅 $Qp(j)$ が大きくなることを意味している。すなわち、動きの大きいシーンほど量子化幅は大きくなり、逆に動きの小さいシーンほど量子化幅は小さくなり、画像は鮮明になる。

## 【0096】

＜符号化パラメータ生成部でのフレームレートと量子化幅の修正＞

符号化パラメータ生成部 2 5 1 では、フレームレートと量子化幅の修正に当たり、まず、式 (1)、式 (2) を用いてフレームレートと量子化幅を決める際、上述のシーンの分類の処理で得られたシーンの分類結果 (シーンを構成するフレームのタイプ) を用いて式 (1) に重みパラメータ  $w_{FR}$  を、式 (2) に重みパラメータ  $w_{QP}$  を加え、フレームレートと量子化幅の修正を行う。

【0 0 9 7】

具体的には、フレーム中に動きベクトルがほとんど存在しないタイプ [1] の場合 (図 6 (a) の場合) にはフレームレートを下げて、量子化幅を小さく取る ( $w_{FR}$ ,  $w_{QP}$  とともに小さくする)。

【0 0 9 8】

図 6 (b) の如きタイプ [2] では、カメラの動きが不自然にならないようなるべくフレームレートを上げ、量子化幅は大きくする ( $w_{FR}$ ,  $w_{QP}$  とともに大きくする)。

【0 0 9 9】

図 6 (c) の如きタイプ [3] では、動いているオブジェクトの動き、すなわち動きベクトルの大きさが大きい場合にはフレームレートを修正する ( $w_{FR}$  を大きくする)。

【0 1 0 0】

図 6 (d) の如きタイプ [4] では、ズームの際にほとんどオブジェクトについては注目されていないと思われることから、量子化幅は大きく取り、フレームレートをできる限り上げる ( $w_{FR}$  を大きくし、 $w_{QP}$  も大きくする)。

【0 1 0 1】

図 6 (e) の如きタイプ [5] もフレームレートを上げ、量子化幅を大きくとる ( $w_{FR}$ ,  $w_{QP}$  とともに大きくする)。

【0 1 0 2】

このようにして設定された重みパラメータ  $w_{FR}$ ,  $w_{QP}$  をそれぞれ加えることにより、フレームレートと量子化幅の調整を行う。

【0 1 0 3】

符号化パラメータ生成部 2 5 1 でのフレームレートおよび量子化幅修正処理の

は以上の如きである。

【0104】

画質を維持するための仕組みとして、符号化パラメータ生成部251では、ユーザ指定によるマクロブロック単位での量子化幅変更が可能である（(iv) マクロブロック毎の量子化幅の設定処理）。マクロブロック単位で量子化幅を可変させることが出来るわけである。その処理の詳細を説明する。

【0105】

＜符号化パラメータ生成部でのマクロブロック毎の量子化幅の設定＞

本発明システムにおいては、符号化パラメータ生成部251は、マクロブロック毎の量子化幅の変更指示を受けるとマクロブロック単位で量子化幅を可変するように機能させることができる。

【0106】

MPEG-4 などでは、画像を16×16画素のブロックに分けて、このブロック単位で処理を進めるが、このブロック単位をマクロブロックと呼ぶ。符号化パラメータ生成部251では、ユーザからマクロブロック毎に量子化幅を変化するように指定された場合には、フレーム中にモスキートノイズがで出易いと判定されたマクロブロックやテロップ文字のように、強いエッジが存在すると判定されたマクロブロックに対して、他のマクロブロックよりも量子化幅を小さく設定することで画質改善を図ることもできる。

【0107】

符号化対象フレームに対して、図7のようにマクロブロックMB<sub>m</sub>をさらに4つに分けた小ブロック毎に輝度値の分散を計算する。このとき、輝度値の分散が大きい小ブロック（b<sub>2</sub>）と分散の小さい小ブロック（b<sub>1</sub>，b<sub>3</sub>）が隣り合う場合に、量子化幅が大きいと、そのマクロブロックMB<sub>m</sub>ではモスキートノイズが発生し易い。つまり、マクロブロックMB<sub>m</sub>内でテクスチャが複雑な部分にテクスチャの平坦な部分が隣接するような場合に、モスキートノイズが出易くなる。

【0108】

そこで、輝度値の分散が大きい小ブロックに分散が小さい小ブロックが隣接し

ている場合をマクロブロック毎に判定し、モスキートノイズが出易いと判定されたマクロブロックについては、他のマクロブロックよりも相対的に量子化幅を小さくする。逆に、テクスチャが平坦でモスキートノイズが出にくいと判定されたマクロブロックに対しては他のマクロブロックよりも相対的に量子化幅を大きくし、発生符号量の増加を防ぐようにする。

【0109】

例えば、j 番目のフレーム内の m 番目のマクロブロックについて、マクロブロック内に小さな 4 つのブロックがあるとき、図 7 に示すように

(ブロック k の分散)  $\geq$  MBVarTre1

かつ

(ブロック k に隣接するブロックの分散)  $<$  MBVarThre2

…式 (3)

という組み合わせを満たす小さなブロックがあるならば、この m 番目のマクロブロックをモスキートノイズが出易いマクロブロックであると判定する (MBVarThre1、MBVarThre2 はユーザが定義する閾値)。このような m 番目のマクロブロックに対して

$$Q_p(j)_{-m} = Q_p(j) - q_1 \quad \dots \text{式 (4)}$$

のようにマクロブロックの量子化幅  $Q_p(j)_{-m}$  を小さくする。これに対してモスキートノイズが出にくいと判定された m' 番目のマクロブロックに対しては

$$Q_p(C)_{-m} = Q_p(C) + q_2 \quad \dots \text{式 (5)}$$

のように、マクロブロックの量子化幅  $Q_p(C)_{-m'}$  を上げることで、符号量の増加を防ぐ ( $q_1$ ,  $q_2$  は正の数で  $Q_p(C) - q_1 \geq$  (量子化幅の最小値)、

$Q p 0) + q 2 \leq (\text{量子化幅の最大値})$  を満たす)。

【0 1 1 0】

その際、前述したカメラパラメータの判定で図 6 (b) の平行移動シーン、図 6 (d) のカメラズームのシーンと判定されたシーンについては、カメラの動きに支配されるために画像中のオブジェクトに対する視覚的注目度が低いと思われることから  $q 1$ ,  $q 2$  は小さくとる。

【0 1 1 1】

逆に、図 6 (a) の静止シーン、図 6 (c) の動いている部分が集中しているシーンでは、画像中のオブジェクトに対する視覚的注目度が高いと思われることから  $q 1$ ,  $q 2$  は大きくとる。

【0 1 1 2】

また、文字のようなエッジが存在するマクロブロックについても、量子化幅を小さくすることで文字の部分を明瞭にさせることもできる。フレームの輝度値データに対してエッジ強調フィルタを施し、マクロブロック毎に濃淡値の勾配が強い画素を調べる。画素の位置を集計し、勾配の大きい画素が部分的に集中しているブロックをエッジが存在するマクロブロックであると判断し、式 (4) に従いそのブロックについて量子化幅を小さくし、式 (5) によりその他のマクロブロックの量子化幅を大きくする。

【0 1 1 3】

このようにしてマクロブロック単位で量子化幅を変更することで、画質を保証できる仕組みが確保できるようになる。

【0 1 1 4】

以上が、符号化パラメータ生成部 2 5 1 における符号化パラメータ生成に当たって実施する (i) フレームレート算出処理、(ii) 量子化幅算出処理、(iii) フレームレートと量子化幅の修正処理、(iv) マクロブロック毎の量子化幅の設定処理、の 4 種の処理の詳細である。

【0 1 1 5】

次に、このようにして算出された符号化パラメータを、ユーザ指定のビットレートを満たすように修正する符号化パラメータ修正部 2 5 3 での処理について詳

しく説明する。

【0 1 1 6】

＜符号化パラメータ修正部での発生符号量の予測＞

符号化パラメータ修正部 2 5 3 での発生符号量の予測は次のようにして行う。符号化パラメータ生成部 2 5 1 により、上述のように算出された各シーン毎のフレームレートと量子化幅を用いて符号化すると、シーンのビットレートの割合が許容されるビットレートの上限值あるいは下限値を超える場合がある。そのため、限界値を超えるシーンのパラメータについてはそれを調整して上限値あるいは下限値内に収まるようにする必要がある。

【0 1 1 7】

例えば、前記算出された符号化パラメータのフレームレートと量子化幅で符号化し、ユーザが設定したビットレートに対する各シーンのビットレートの割合を算出したとき、図 8 (a) のようにビットレートの上限值あるいは下限値を超えるようなシーン (S 3, S 6, S 7) が出てくる場合がある。

【0 1 1 8】

そこで本発明では、符号化パラメータ修正部 2 5 3 により、次のような処理を実施して、それぞれのシーンのビットレートの割合が、許容されるビットレートの上限值あるいは下限値を超えないように修正する処理を施す。

【0 1 1 9】

すなわち、ユーザが設定したビットレートに対する割合を算出したとき、ビットレートの上限值を超えるようなシーン (S 3, S 6) では図 8 (b) のようにビットレートを上限値に設定し直す。同様に下限値を下回るシーン (S 7) では図 8 (b) のようにビットレートの割合を下限値に設定し直す。

【0 1 2 0】

この操作により、過剰、あるいは不足となった符号量は、図 8 (c) のように修正しなかった他のシーンに再分配し、全体の符号量は変えないように操作する。

【0 1 2 1】

そのためには、発生符号量の予測が必要である。ここでは、発生符号量は例え

ば次のようにして予測する。

【0 1 2 2】

符号化パラメータ修正部 2 5 3 は、各シーンの最初のフレームを I ピクチャ、その他を P ピクチャにすると仮定し、それぞれの符号量を算出する。まず I ピクチャの発生符号量を推定する。I ピクチャの発生符号量については一般的に量子化幅 Q P と符号量の間に、図 9 のような関係があることから、1 フレームあたりの発生符号量 C o d e I を例えば次のように算出する。

【0 1 2 3】

$$C o d e I = I a \times Q P ^ I b + I c \quad \dots \text{式 (6)}$$

ここで I a, I b, I c は画像サイズ等により定められる定数とする。また、<sup>^</sup> はべき乗を表す。

【0 1 2 4】

さらに、P ピクチャについては、動き補償残差 M e S a d と符号量の間に、ほぼ図 1 0 に示す如きの関係があることから、1 フレームあたりの発生符号量 C o d e P を例えば次のように算出する。

【0 1 2 5】

$$C o d e P = P a \times M e S a d + P b \quad \dots \text{式 (7)}$$

ここで、P a, P b は画像サイズ、量子化幅 Q p 等により定められる定数とする。画像特徴量計算部 2 2 0 において、式 (7) に用いる M e S a d は、既に求められているものとし、これらの式から各シーン毎に発生する符号量の割合を算出する。J 番目のシーンの発生符号量は

$$C o d e (j) = C o d e I + (\text{符号化する予定のフレームの } C o d e P \text{ の和})$$

$$\dots \text{式 (8)}$$

上記の式により算出されたシーン毎の符号量  $Code(j)$  を、そのシーンの長さ  $T(j)$  で除算するとそのシーンの平均ビットレート  $BR(j)$  が算出される。

【0126】

$$BR(j) = Code(j) / T(j) \quad \dots \text{式 (9)}$$

このように算出されたビットレートをもとに、符号化パラメータの修正を行う。また、上記のようなビットレートの修正により予測された符号量を大幅に変更するような場合、各シーンのフレームレートを修正してもよい。すなわち、ビットレートを低くしたシーンではフレームレートも低くし、ビットレートを高めたシーンではフレームレートも高めることにより画質を保つようにする。

【0127】

以上が、符号化パラメータ修正部 253 での個々の処理の詳細である。

【0128】

以上のように、本発明は、動画像信号の符号化をするに当たり、状態を把握して調整するための予備処理（第1パス）を実施し、得られた結果を用いて符号化を実施する（第2パス）の2段階処理形態としたものであって、動画像信号についてシーン毎のフレームレートとビットレートを求める第1パスの処理を実施し、当該第1パス目で算出されたシーン毎のフレームレートとビットレートを第2パス目で符号化部に渡し、動画像信号を符号化することにより、フレームスキップや画質の劣化の劣化のない動画像符号化を実施できるようにしたものである。符号化部では、1パス目で得られた符号化パラメータをもとに、シーン毎に目標ビットレートとフレームレートを切り替えながら、従来からのレート制御を用いて符号化する。また、第1パス目で得られたマクロブロックの情報を用いて、レート制御により算出された量子化幅に対して相対的にマクロブロックの量子化幅を変化させる。これにより、まとまったひとつのシーンの中ではビットレートが保たれるので、符号化されたビットストリームのサイズは目標データサイズを満

たすことができるようになる。

【0129】

比較のために、図11に、本発明の手法と従来法を用いて符号化した際のビットレートとフレームレートの推移の例を示しておく。

【0130】

図11の(a)が従来法におけるビットレートとフレームレートの推移の例であり、図11の(b)が本発明手法におけるビットレートとフレームレートの推移の例である。

【0131】

従来法では、図11(a)の〔I〕に示すように、一定の目標ビットレート401が定められており、これに対して403のように一定のフレームレートが設定されている。また、実際のビットレートとフレームレートは図11(b)の〔I〕に示すように、402(実際のビットレート)および404(実際のフレームレート)の如きとなる。このとき、動画像が動きの激しいシーンに切り替わると(t11~t12区間参照)、このような動画像では発生符号量が急増するために、図15(b)のようなフレームスキップが起こり、図11(b)の〔II〕に符号404で示すように、フレームレートが落ちてしまう。

【0132】

これに対して本発明の手法(図11(b))では、シーンに応じて最適な値になるように405の如く目標ビットレートを定めており、また、シーンに応じて最適な値になるように、407の如く目標フレームレートを定めている。

【0133】

これにより、動画像が動きの激しいシーンに切り替わったときは、増大する符号量対応に目標値が変わることから、そのシーンに割り当てられたビットレートが多くなるためにフレームスキップは起こりにくくなり、また、フレームレートは目標値を満たすことができる。

【0134】

(第2の実施形態)

次に、ソースデータがMPEGストリーム(DVDの場合はMPEG-2スト

リーム)である場合に、1パス目にビットストリームを全て再生するのではなく、必要な信号のみを部分的に再生することで1パス目の処理量を削減する例について説明する。

【0135】

ここでの構成例は、第1の実施形態で用いたものと基本的には同じでよい。

【0136】

ソースデータがMPEGストリームの場合、そのビットストリームの構成は図12の如きである。図12に示す例のように、MPEGストリームは、フレーム内符号化／フレーム間符号化等を切り替えたりするためのモード情報と、フレーム間符号化する際の動きベクトル情報と、輝度や色差信号を再生するためのテキストチャ情報に大別される。

【0137】

ここで、モード情報によりフレーム内符号化するブロックの数が多い場合には、シーンチェンジが発生していると想定されるため、特徴量計算部220(図1参照)でのシーン区切りの判定に利用できる。

【0138】

また、MPEGストリームには動きベクトル情報が含まれているので、このMPEGストリーム中の動きベクトル情報を抽出して特徴量計算部220で利用すれば良い。

【0139】

すなわち、特徴量計算部220では、動画像信号のシーン分割、動画像信号の各フレームにおける画像特徴量(動きベクトルの数、分布、ノルムの大きさ、動き補償後の残差、輝度・色差の分散等)を求める処理をするわけであるが、第1の実施形態のように、これら全てを演算処理により求めるのではなく、ここでは、MPEGストリーム中の“モード”の情報により、フレーム内符号化するブロックの数の大小を知り、これよりシーン区切りの判定してシーン分割処理に代え、また、MPEGストリームの“動きベクトル”の情報を抽出してそのまま流用することで動きベクトル演算処理を省略する。

【0140】

このように、MPEGストリームは、全てのデータを再生せずとも、一部の情報を再生するだけで特徴量計算部 2 2 0 で利用できるデータを、MPEGストリーム中から取得できることを利用して処理を簡素化することができる。

【0 1 4 1】

このような部分的に再生された信号を利用する場合は、図 1 の構成において、デコーダ 2 1 0 によりデコードされた信号から上記“モード”の情報と、“動きベクトル”の情報を得てこれを信号線 2 7 を介して特徴量計算部 2 2 0 に供給する構成とし、特徴量計算部 2 2 0 には、動画像信号のシーン分割、動画像信号の各フレームにおける画像特徴量の演算処理のうち、“モード”の情報をを用いたフレーム内符号化するブロックの数の大小からシーン区切りの判定してシーン分割処理を行う構成とし、MPEGストリームの“動きベクトル”の情報をそのまま流用して動きベクトルの数を取得する構成とし、他の演算（動きベクトルの分布、ノルムの大きさ、動き補償後の残差、輝度・色差の分散等）については第 1 の実施の形態と同様の処理をする構成とする。

【0 1 4 2】

この構成により、特徴量計算部 2 2 0 の処理は一部を簡略化した構成として実現できることになる。

【0 1 4 3】

なお、本発明において、上記実施形態には種々の段階の発明が含まれており、開示される複数の構成要件における適宜な組み合わせにより種々の発明が抽出され得る。例えば、実施形態に示される全構成要件から幾つかの構成要件が削除されても、発明が解決しようとする課題の欄で述べた課題の少なくとも 1 つが解決でき、発明の効果の欄で述べられている効果の少なくとも 1 つが得られる場合には、この構成要件が削除された構成が発明として抽出され得る。

【0 1 4 4】

また、本発明における実施形態に記載した手法は、コンピュータに実行させることのできるプログラムとして、磁気ディスク（フレキシブルディスク、ハードディスクなど）、光ディスク（CD-ROM、CD-R、CD-RW、DVD、MOなど）、半導体メモリなどの記録媒体に格納して頒布することもでき、また

、ネットワークを介しての伝送により、頒布することもできる。

【0 1 4 5】

【発明の効果】

以上記述した如く、本発明によれば、動画像を解析し、シーン内容を編集操作に活用すると共に、編集操作により作成された新たな動画像に対してもシーン毎の統計的特徴量の相対的關係から最適な符号化パラメータが算出されるため、編集操作が容易になると共に、シーン毎にまとまりのある画像が得られ画質改善効果が得られる。

【図面の簡単な説明】

【図 1】

本発明を説明するための図であって、本発明の一実施形態に係る動画像符号化装置の構成を示すブロック図である。

【図 2】

本発明を説明するための図であって、構造化情報提示手段の表示例を説明する図である。

【図 3】

本発明を説明するための図であって、符号化するシーンを部分的に選択する場合の説明図である。

【図 4】

本発明を説明するための図であって、本発明システムにおける最適パラメータ計算部の構成例を示すブロック図である。

【図 5】

本発明を説明するための図であって、本発明の実施形態におけるシーン分割の処理手順の例を示すフローチャートである。

【図 6】

本発明を説明するための図であって、本発明の実施形態における動きベクトルによるフレームのタイプ分けについて説明する図である。

【図 7】

本発明を説明するための図であって、本発明システムにおけるモスキートノ

イズの発生し易いマクロブロックの判定について説明する図である。

【図 8】

本発明を説明するための図であって、本発明システムにおける発生符号量調整の処理手順を示す図である。

【図 9】

本発明を説明するための図であって、本発明システムにおける I ピクチャに関する発生符号量の推移を示す図である。

【図 10】

本発明を説明するための図であって、本発明システムにおける P ピクチャに関する発生符号量の推移を示す図である。

【図 11】

本発明システムにおけるビットレートとフレームレートの推移を従来法と比較した図である。

【図 12】

MPEGビットストリームの例である。

【図 13】

検証モデルのエンコーダブロック図である。

【図 14】

従来のレート制御について示す図である。

【図 15】

従来のレート制御を説明する図である。

【符号の説明】

- 1 0 0 …エンコーダ
- 1 2 0 …サイズ変換部
- 2 0 0 …ソースデータ
- 2 1 0 …デコーダ
- 2 2 0 …特徴量計算部
- 2 3 0 …構造化情報蓄積部
- 2 4 0 …構造化情報提示装置

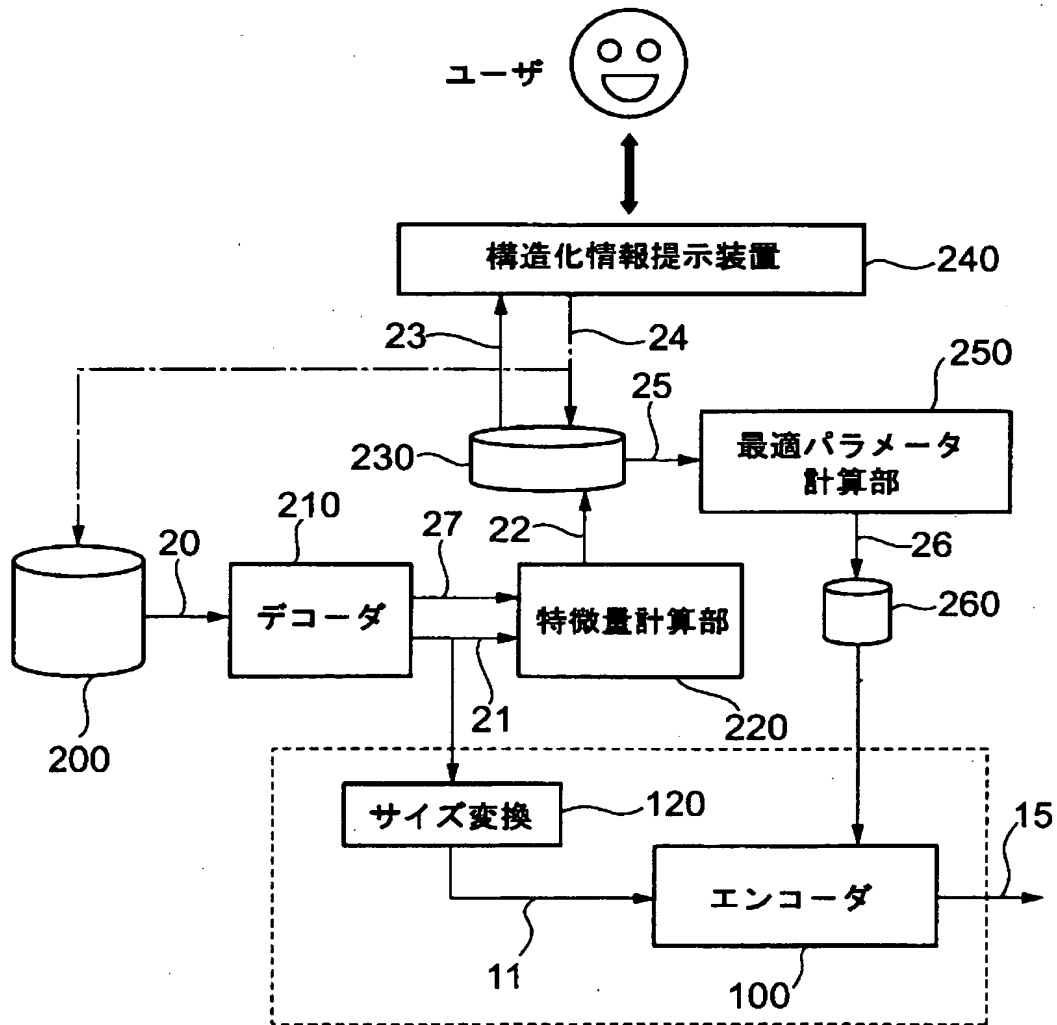
- 2 5 0 …最適パラメータ計算部
- 2 5 1 …符号化パラメータ生成部
- 2 5 2 …発生符号量予測部
- 2 5 3 …符号化パラメータ修正部
- 2 6 0 …最適パラメータ蓄積部。

【書類名】

図面

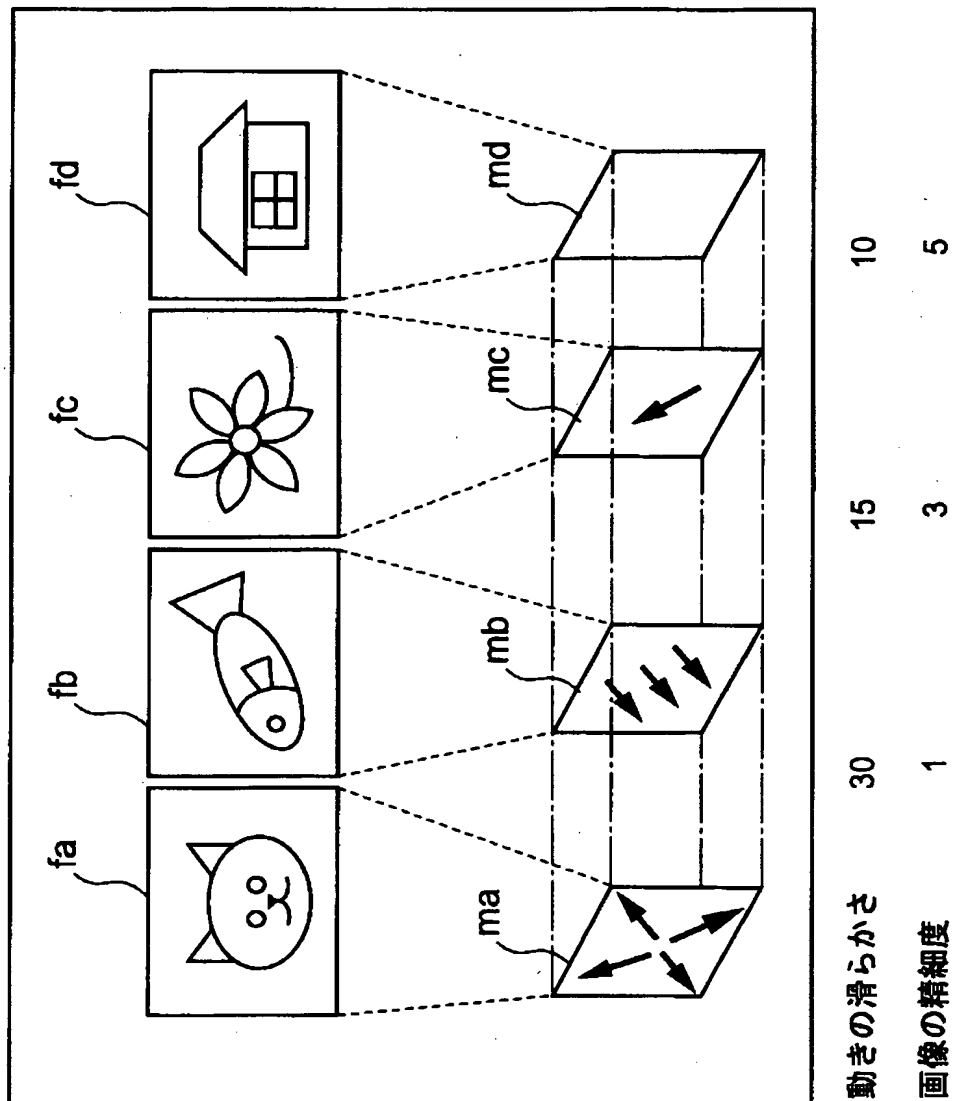
【図 1】

最適符号化処理の流れ図



【図 2】

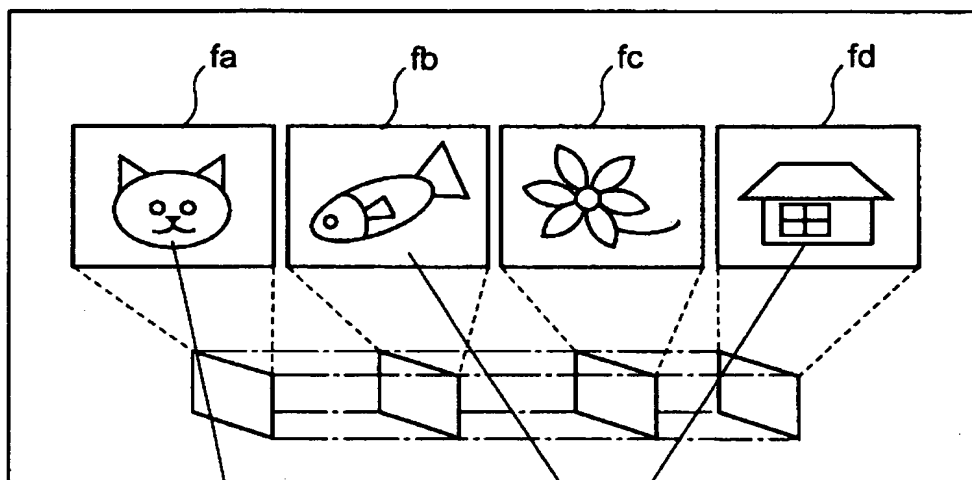
構造化情報提示手段の例



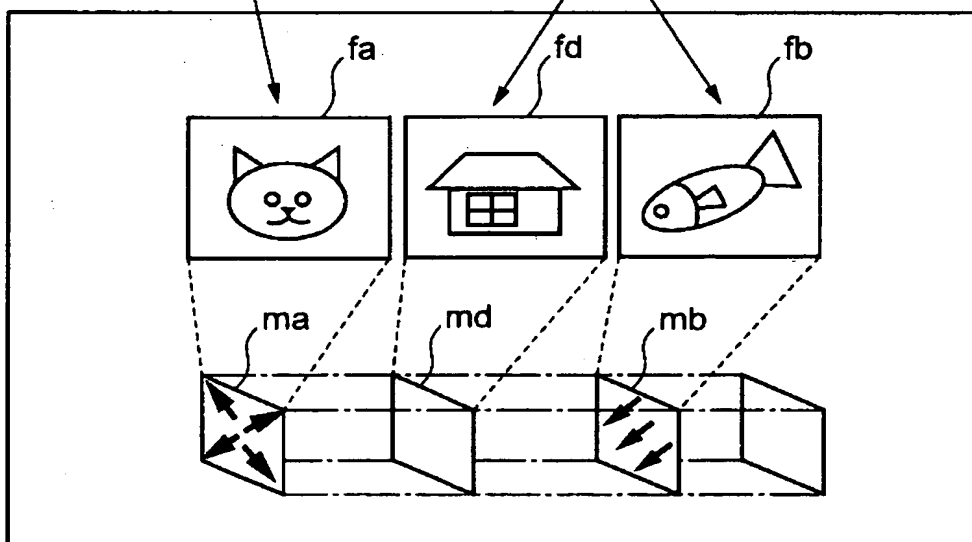
【図 3】

符号化するシーンを選択する例

(a)

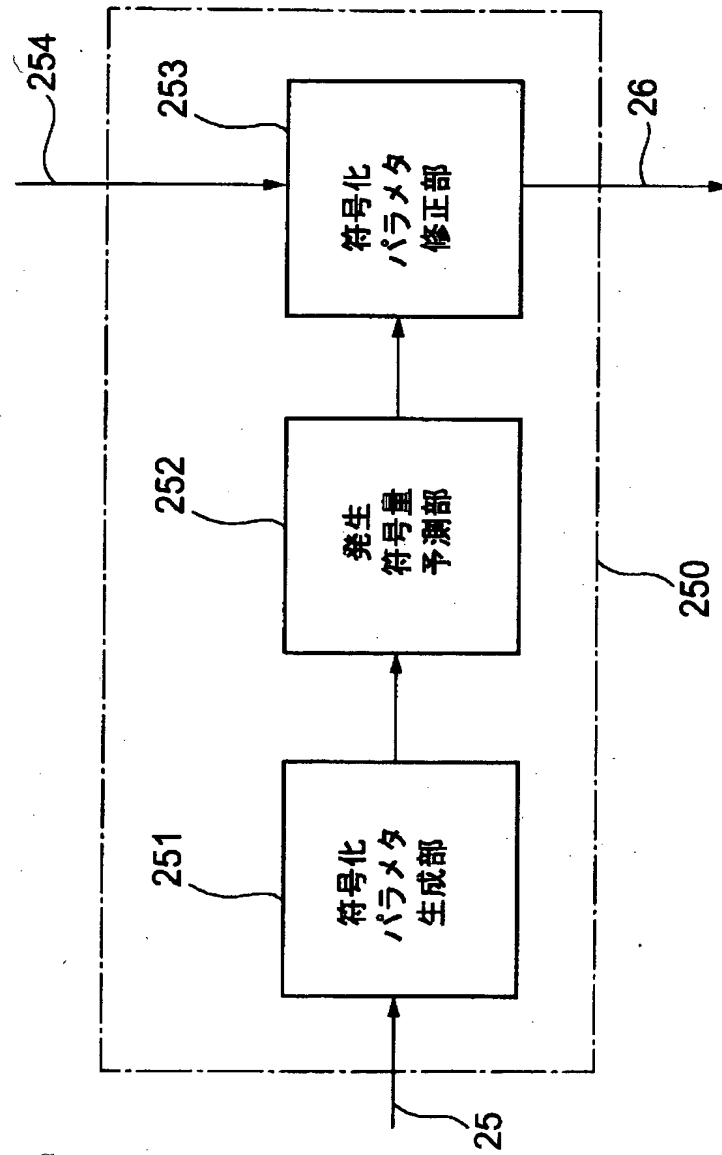


(b)



【図 4】

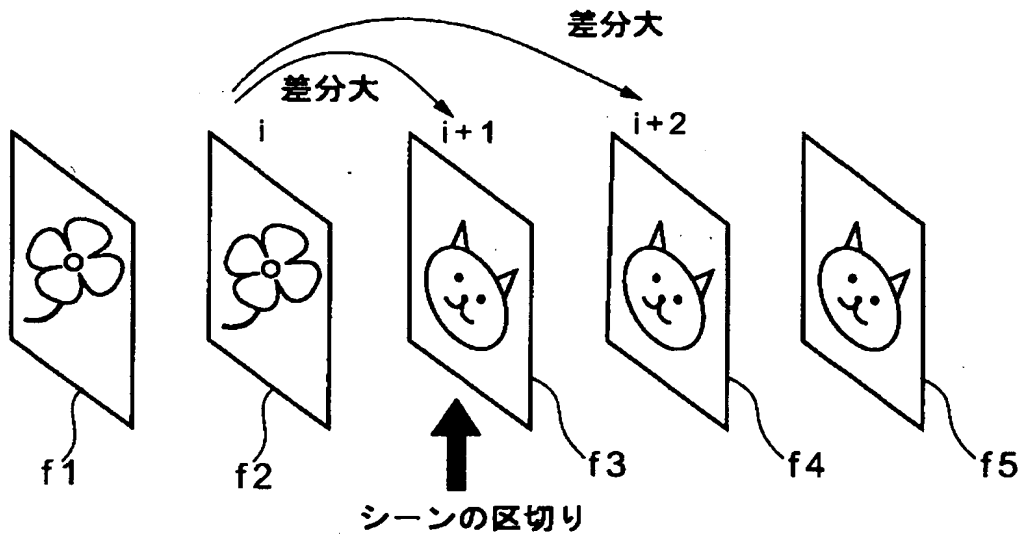
最適パラメタ計算部の説明図



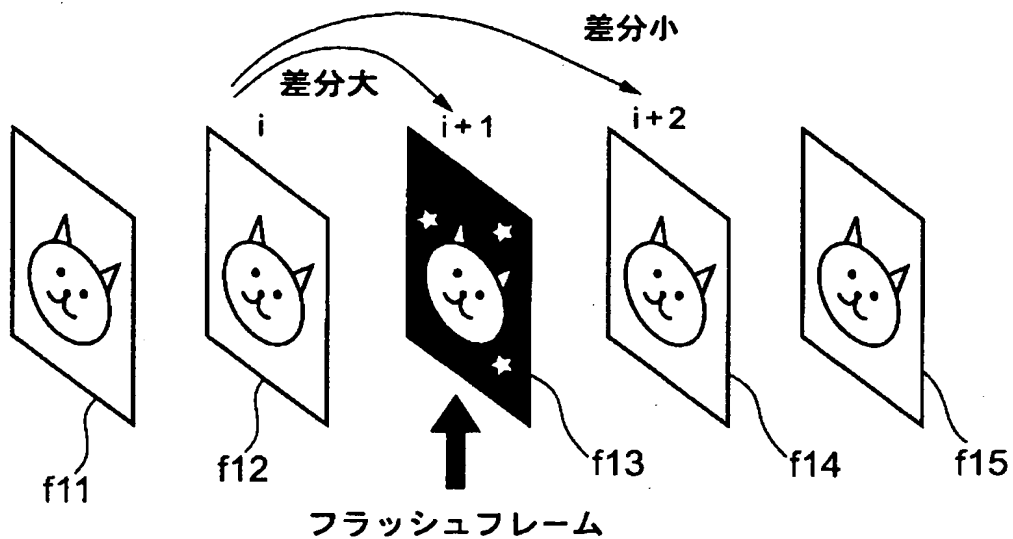
【図 5】

# シーン区切りの判定

## (a) シーン区切りの場合



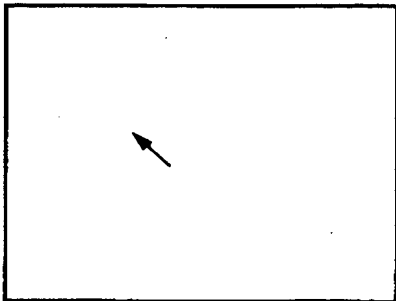
## (b) フラッシュフレームの場合



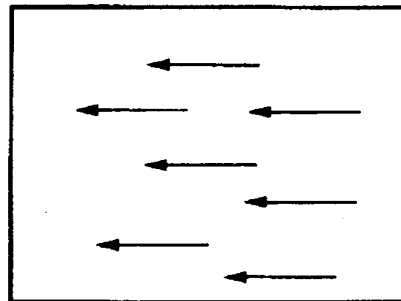
【図 6】

# 動きベクトルの分類

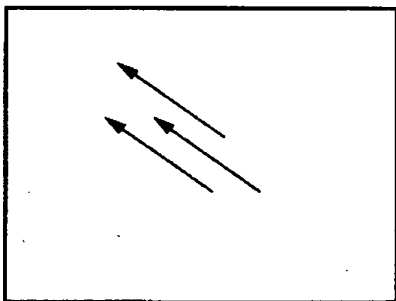
(a)動きベクトルがほとんどない場合



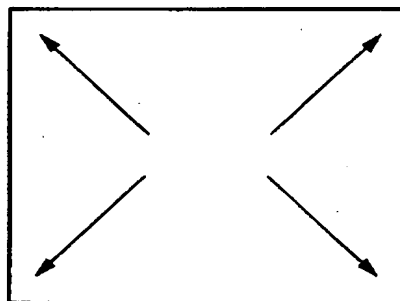
(b)同じ向き・大きさの動きベクトルが画面全体に分布している場合



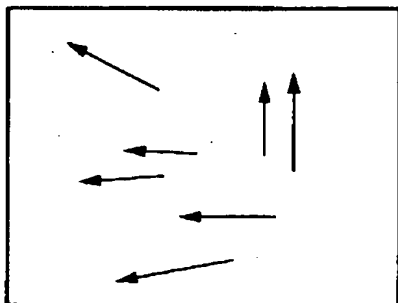
(c)フレーム中で特にある部分だけ動きベクトルが現れている場合



(d)フレーム中に放射状に動きベクトルが分布している場合

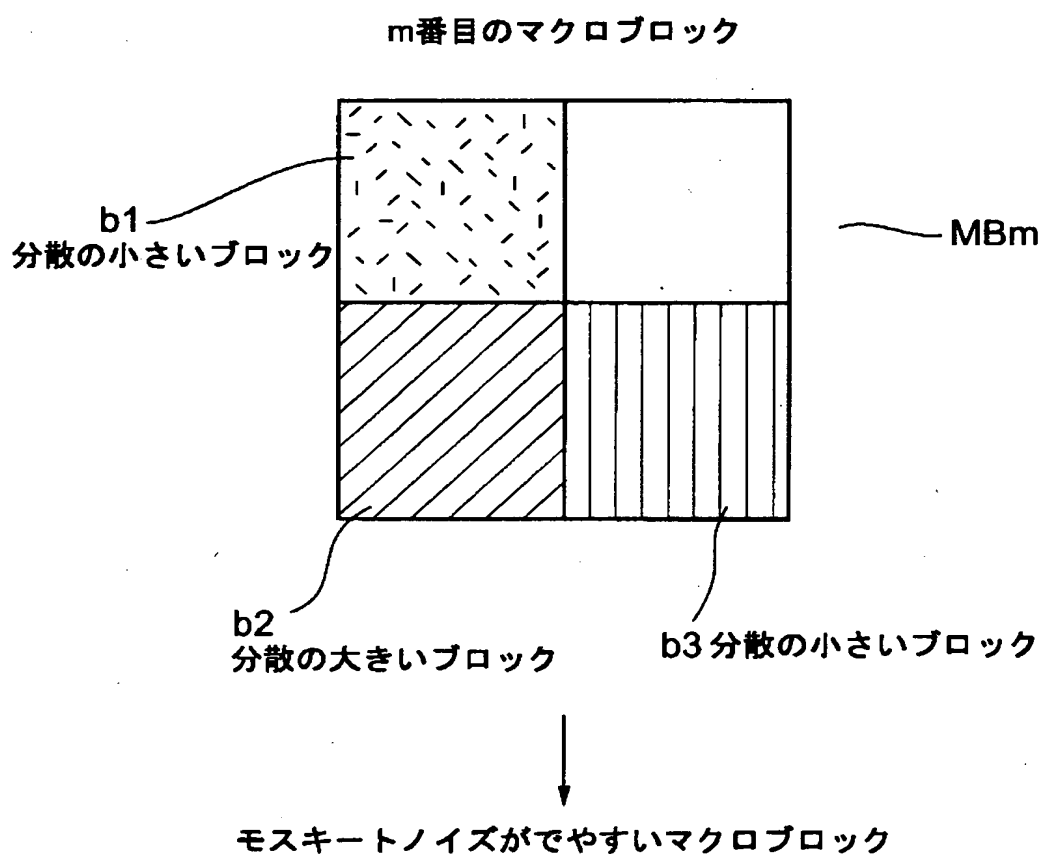


(e)フレーム中の動きベクトルの数が多く、方向もばらばらである場合



【図 7】

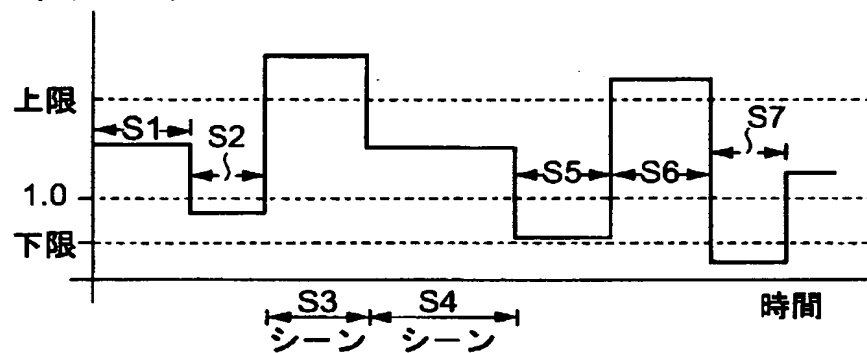
モスキートノイズがしやすいマクロブロックの判定



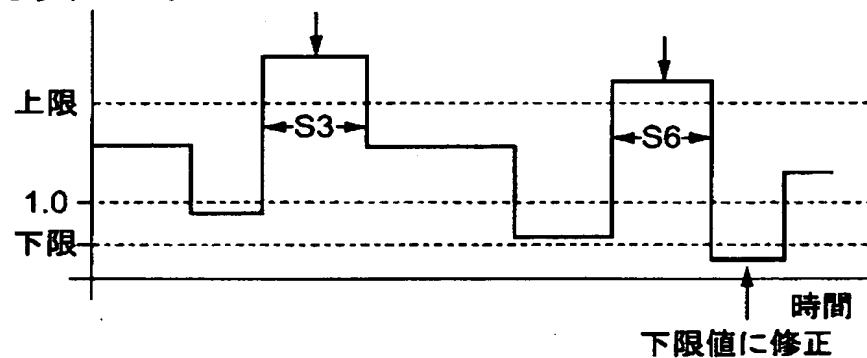
【図 8】

# ビットレートの調整

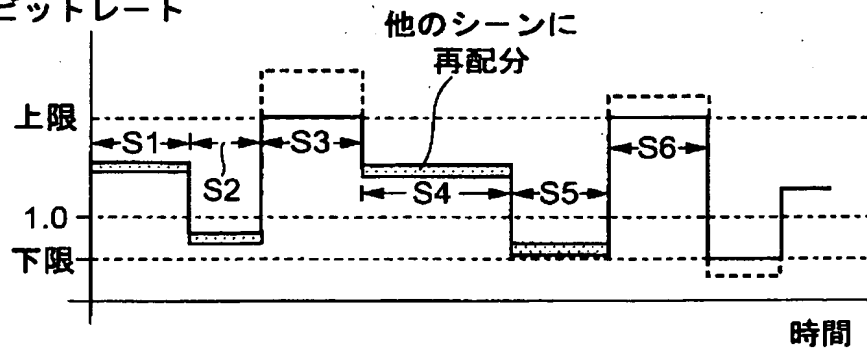
(a) シーンの平均  
ビットレート



(b) シーンの平均  
ビットレート 上限値に修正

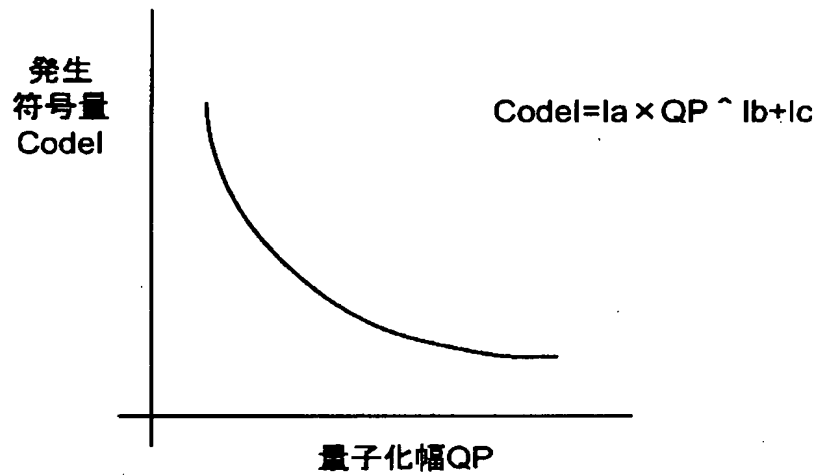


(c) シーンの平均  
ビットレート



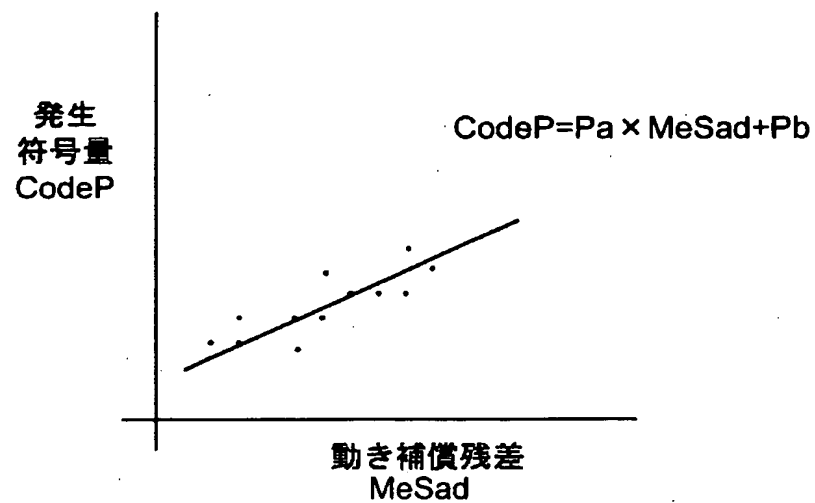
【図 9】

I ピクチャに関する発生符号量



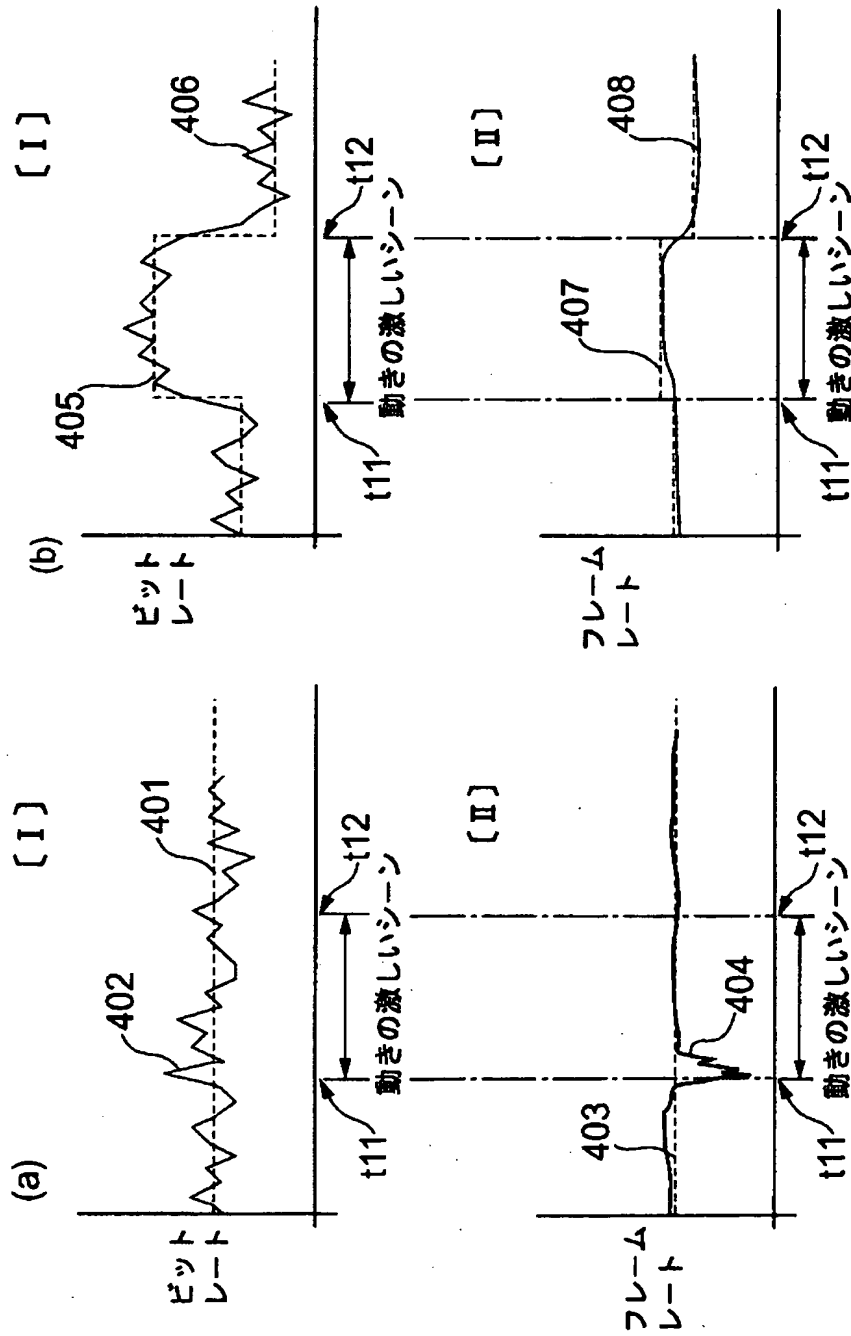
【図 10】

P ピクチャに関する発生符号量



【図 1 1】

従来法と本発明による符号化画像の  
ビットレートとフレームレートの推移



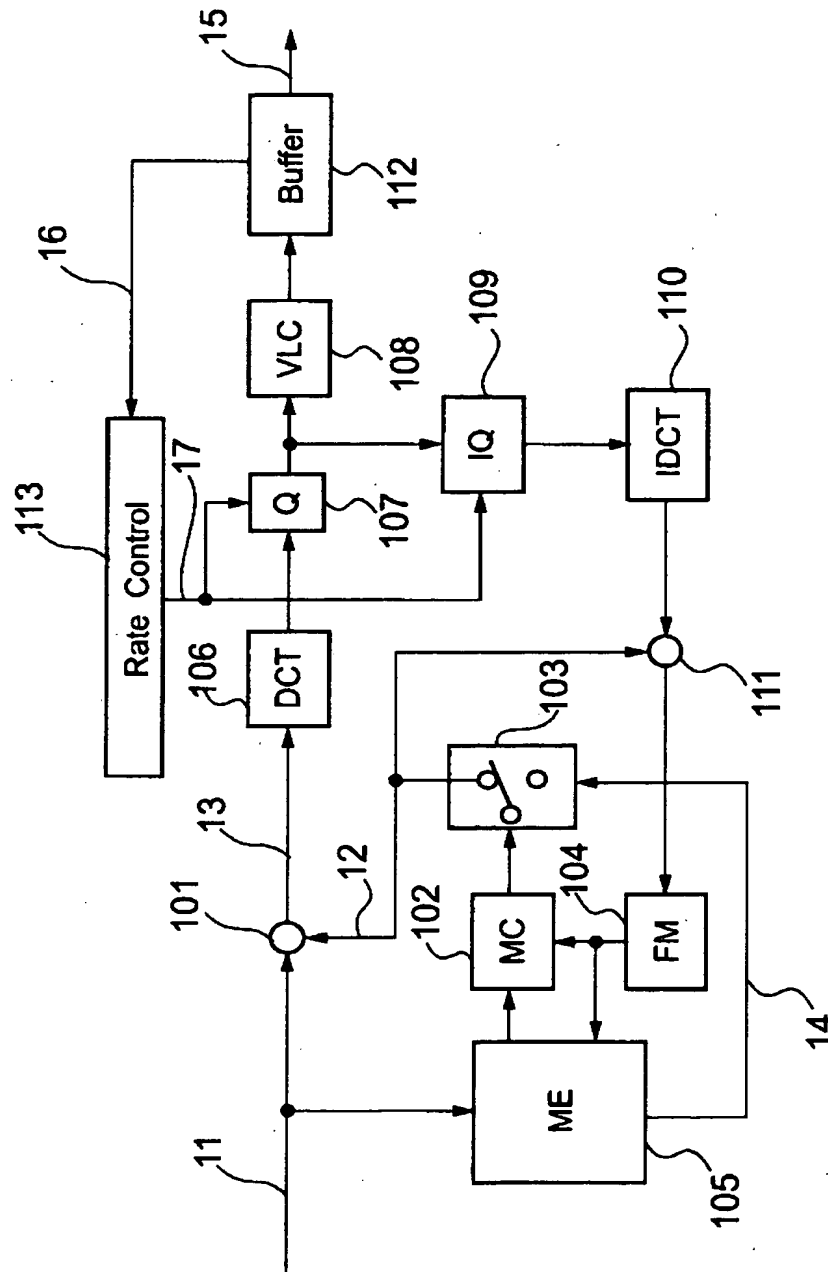
401...目標ビットレート 403...目標フレームレート 405...目標ビットレート 407...目標フレームレート  
402...実際のビットレート 404...実際のフレームレート 406...実際のビットレート 408...実際のフレームレート

【図 1 2】

MPEGビットストリームの例

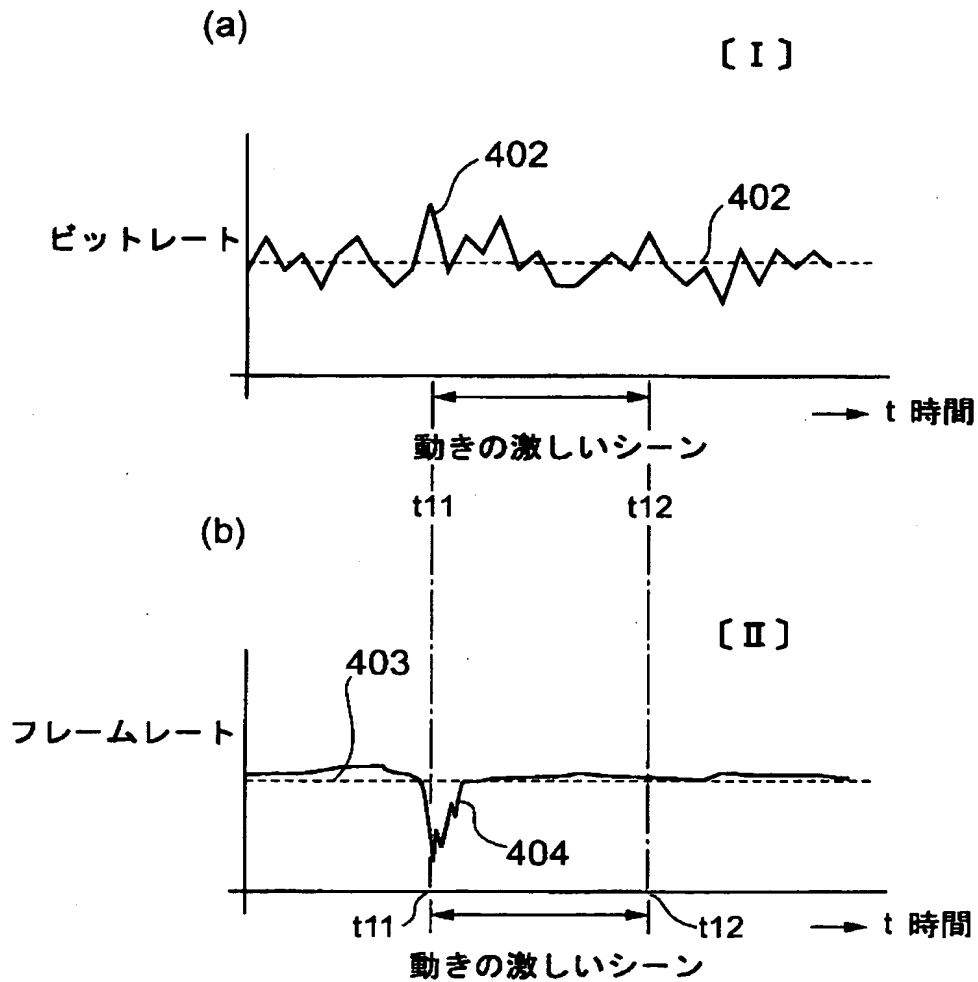
モード	動きベクトル	テクスチャ
-----	--------	-------

【図 1 3】



【図 1 4】

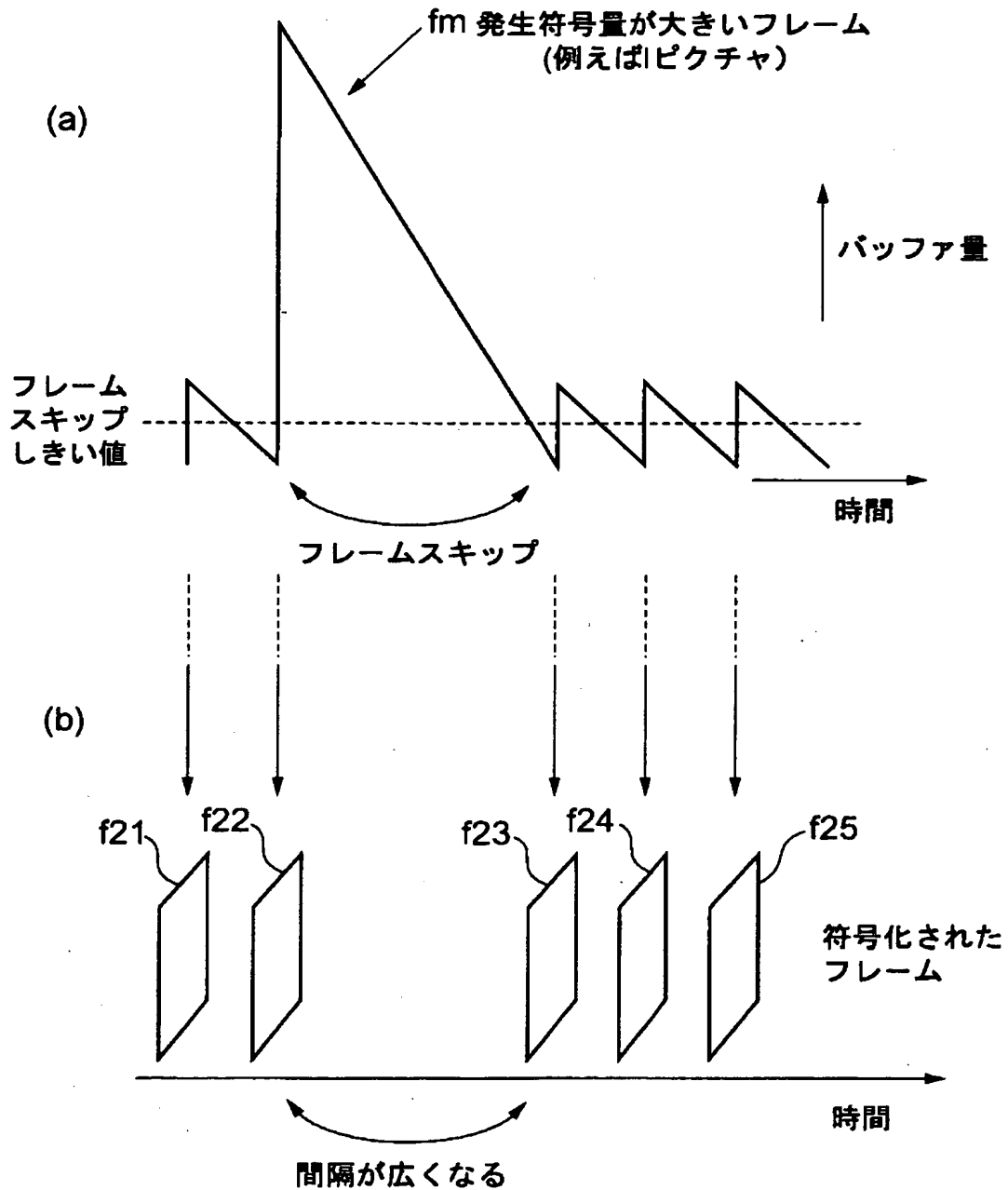
従来法による符号化画像の  
ビットレートとフレームレートの推移



- 401…目標ビットレート
- 402…実際のビットレート
- 403…目標フレームレート
- 404…実際のフレームレート

【図 1 5】

従来のバッファとフレームレートの関係



【書類名】 要約書

【要約】

【課題】 動画像符号化においてオブジェクトの動きが激しいシーンにおいて急激にフレームレートが落ちたり、量子化幅が適切でないために画像が歪むなど、画質劣化が目立ち易いという問題点を解決する。

【解決手段】 動画像を解析して、フレーム(FM)毎に統計的特徴量を算出する第1特徴量算出手段220と、この特徴量算出手段にて算出された特徴量に応じて動画像を時間的に連続した少なくとも一つのFMからなるシーン(SC)に分割するシーン分割手段220と、シーン分割手段により分割された各SC毎の平均的な特徴量を第1特徴量算出手段より得られたFM毎の特徴量から求める第2特徴量算出手段220と、シーン分割手段により分割された複数SCの一部あるいは全てを選択するシーン選択手段240と、シーン選択手段により選択されたSC毎の特徴量からSC毎の符号化パラメータを生成するパラメータ生成手段250と、この生成された符号化パラメータを用いて入力動画像信号を符号化する手段100とを具備する。

【選択図】 図1

出 願 人 履 歴 情 報

識別番号 [000003078]

1. 変更年月日 1990年 8月22日  
[変更理由] 新規登録  
住 所 神奈川県川崎市幸区堀川町72番地  
氏 名 株式会社東芝
2. 変更年月日 2001年 7月 2日  
[変更理由] 住所変更  
住 所 東京都港区芝浦一丁目1番1号  
氏 名 株式会社東芝